Distribution of Matter in the Universe: from Lensing Clusters to Large Scale Structure

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Irshad Mohammed

von

Indien

Promotionskomitee

Prof. Dr. Romain Teyssier (Vorsitz)
Dr. Prasenjit Saha
Prof. Dr. Ben Moore

Zürich, 2015

Irshad Mohammed
Physik Institute and Institute for Computational Science
University of Zurich
Winterthurerstrasse 190
CH-8057 Zürich
Switzerland
irshad@physik.uzh.ch

Contents

To My Family

Acknowledgments

I would like to express deepest gratitude to my committee chair, Dr. Prasenjit Saha, who has the attitude and substance of a genius. He continously and convincingly conveyed a spirit of adventure with regard to research, and an excitement in teaching. Without his guidance and persistent help, this dissertation would not have been possible.

I would like to thank my joint advisors, Professor Uroš Seljak and Professor Romain Teyssier, whose work demonstrated to me the excellence of scientific research, and a need of critical thinking about the future of cosmology and astrophysics.

I would like to thank Janu Verma, a very dear friend, for reading the manuscript, and providing necessary stylistic improvements. I would also like to thank Christian Reinhardt for carefully reading the dissertation, and making useful comments.

In addition, I would like to thank fellow students and post-doctoral associates in the institute for fruitful conversations about science, and life in general. I would like to thank everyone in the *Physik Institute* and the *Institute for Computational Science* for establishing an ideal environment to carry out scientific research.

Last but not the least, I would like to thank my family and friends for keeping things easy around me, so I could keep working without difficult moments.

Abstract

The distribution of matter in the Universe contains a wealth of information about the energy content in the Universe, its properties, and evolution. It can be studied in two very different regimes. First, in gravitationally bound systems like galaxies, cluster of galaxies etc.; second, in the large scale structure (LSS) of the Universe. Each of these regimes have specific applications and they collectively improve the understanding of the theory of structure formation and cosmology. Firstly, clusters of galaxies are the largest gravitationally bound structures in the Universe, consisting of hundreds of galaxies and intra-cluster gas moving in the potential well of the large dark-matter component. Becasue of their deep potential well, high density and high temperature of their gas, the clusters can be studied with probes like gravitational lensing, X-ray observations etc., and provide cosmic laboratories to study the interactions of baryons and dark-matter, or non-standard properties of dark-matter, if any. Secondly, the LSS is formed due to the evolution of tiny perturbations in the initial density field via gravitational instability and many baryonic processes. By studying the distribution of matter in the LSS, it is possible to constrain the initial conditions and/or the cosmological parameters.

In this PhD dissertation, I focussed on these two aspects and successfully concluded five scientific papers, which are attached in this manuscript:

We studied the distribution of matter in three clusters of galaxies and reconstructed their mass distribution using a non-parametric technique in strong gravitational lensing. In two of these clusters, we found significant offset between the density peaks and the nearest galaxy. We discussed weather these offsets could have an astrophysical origin or be an indication of self-interactions of dark-matter particles. Continuing in the same vein, we studied the effect on time delay, between different images of the same source, of the mass distribution of the lensing clusters. We found that in clusters where the steepness degeneracy is already broken by multiple background sources at different redshifts, time delay information can be used to constrain the lopsidedness of the cluster core.

In other work, we built an analytical model for the matter power spectrum that describes the matter density fluctuations statistically (only to second order). The model is computationally inexpensive and predicts the matter power spectrum to a percent level accuracy up to $k \sim 0.7h^{-1}$Mpc.

Furthermore, we studied the effects of baryons on the sky-projected weak lensing shear power spectrum. We argued that these effects become significant at small scales $\ell \sim 5000$ and if ignored, it will bias the interpretation of the cosmological parameters to many sigma.

Finally, we reconstructed the mass maps of six Hubble Frontier Field clusters. Their mass distribution shows elongation, multiple-cores, and many sub-structures indicating a recent major merger. We also quantified their clustering properties with the power spectrum of the mass field and compared them with ΛCDM simulated clusters.

Zusammenfassung

Die Verteilung der Materie im Universum enthält eine Fülle von Informationen über die Eigentschaften, den Energieinhalt und die zeitliche Entwicklung des Universums. Sie kann in zwei sehr unterschiedlichen Grössenordnungen untersucht werden, einerseits in gravitativ gebundenen Systemen oder Subsystemen wie zum Beispiel Galaxien und Galaxienhaufen und andererseits in den grössten Strukturen (Large Scale Structures, LSS) des Universums. In jedem dieser beiden Regime finden man spezifische Anwendungen und zusammen bilden sie ein Werkzeug für das Verständnis der Strukturbildung und der Kosmologie im Allgemeinen. Galaxienhaufen sind die grössten gravitativ gebundenen Strukturen im Universum und bestehen von Hunderten von Galaxien. Das Gas im Inneren bewegt sich in der Potentialtopf welche von der dunklen Materiekomponenten des Systems erzeug wird. Wegen der Tiefe des Potentialtopfs und der hohen Gasdichte und Temperatur kann ein solcher Haufen mit Techniken wie Gravitationslinsen, Röntgenbeobachtungen usw untersucht werden. Es bietet eine Art kosmisches Labor, in dem die Wechselwirkungen zwischen Baryonen und dunkler Materie zu studiert oder Abweichungen der dunklen Materie vom Standartmodell untersucht werden können. Die Grössstrukturen andererseits entstehen aus dem gravitativen Kollaps winziger Fluktuationen im ursprünglichen Dichtefeld. Das Studium der Materieverteilung in LSS ermöglicht es, die Anfangsbedingungen und/oder kosmologischen Parameter einzuschränken.

Im Rahmen dieser Doktorarbeit habe ich mich mit diesen beiden Aspekten auseinandergestetzt und erfolgreich fünf wissenschaftliche Arbeiten publiziert welche in dieser Doktorarbeit in gebundener Form vorliegen.

Als erstes untersuchten wir die Verteilung der Materie in drei Galaxienhaufen und rekonstruierten ihre Massenverteilungen mit Hilfe einer parameterfreien Formulierung des starken Gravitationslinseneffekts. In zwei dieser Haufen fanden wir signifikante Abweichungen zwischen den Dichtemaxima der Messungen und der nächsten Galaxie. Wir untersuchten, ob diese Abweichungen einen astrophysikalischen Ursprung haben oder Hinweise auf Selbstinteraktionen zwischen dunkler Materie sind.

Als nächstes haben wir — einer ähnlichen Richtung folgend — die Auswirkungen der Massenverteilung des beobachteten Haufens auf die Zeitverzögerung zwischen den verschiedenen Bildern der selben Quelle untersucht. Wir fanden, dass in Galaxienhaufen die Zeitverzögerung zwischen Bildern Informationen liefert, welche verwendet werden können um die Kugelsymmetrie des Kerns des Haufens zu beschränken.

In den anderen Arbeiten haben wir ein analytisches Modell für das Materieleistungsspektrum entwickelt, das die Materiedichtefluktuationen statistisch bis zur zweiten Ordnung beschreibt. Das Modell ist rechnerisch kostengünstig und kann das Materieleistungsspektrum bis auf einen Prozent Genauigkeit von $k \sim 0.7h^{-1}\mathrm{Mpc}$ voraussagen.

Darüber hinaus untersuchten wir die Auswirkungen der Baryonen auf das

schwache Linsen Scherleistungsspektrum. Wir argumentierten, dass diese Effekte auf kleinen Skalen $\ell \sim 5000$ signifikant werden und bei Vernachlässigung einen systematischen Fehler in die Messung und Interpretation der kosmologischen Parameter von einigen Sigma Abweichung einführt.

Schliesslich rekonstruierten wir die Massenkarten von sechs Hubble Frontier Field Haufen. Ihre Massenverteilung zeigt Dehnung, mehrere Kerne und viele Substrukturen, welche einen Hinweis auf eine kürzliche grosse Fusion liefern. Wir messen auch deren Klumpigkeit mit dem Leistungsspektrum des Massenfeldes und vergleichen sie mit Simulationen welche dem ΛCDM Model folgen.

"There are only certain intervals of time when life of any sort is possible in an expanding universe and we can practise astronomy only during that habitable time interval in cosmic history."

– *John D. Barrow*

1

Introduction

Our understanding of the Universe has advanced significantly in last two decades with the development in our theoretical, observational, and computational abilities. While we are confident about the flatness of the Universe from Cosmic Microwave Background (CMB) experiments, the indicative accelerating Universe from Supernovae-Ia surveys which is well supported by a cosmological constant. Finally, a non-interacting, collision-less dark matter is essential to explain the kinematics of galaxies within clusters, light-curves of galaxies, and gravitational lensing etc.. A comprehensive model, called ΛCDM, is very successful in explaining most of the observations independently and their combined constraints tells us that our Universe is composed of about 70% dark-energy in the form of cosmological constant Λ, about 25% non-interacting cold dark-matter (CDM) and only 5% of the ordinary (baryonic) matter [60].

1.1. An expanding Universe

1.1.1. Cosmological principle

Most of the observables in cosmology are statistical in nature. The fact that we have only one Universe to observe, makes it more interesting albeit challenging to test the physical laws. Also because of finite speed of light, we can only observe the current state of the Universe locally. This also gives us the ability to look back into the past stages of the Universe. We cannot observe the past of the Milky Way, we can however observe similar galaxies and draw a typical evolutionary picture. Studying the distribution of galaxies is interesting in many other ways too. The distribution of galaxies clearly indicates that the Universe is highly inhomogeneous around us, but if we go at sufficiently large scales, galaxy distribution is very isotropic [2, 3]. Observations of the CMB also indicates that the Universe is isotropic up to 1 part in 10^5 [59]. If we combine the isotropic principle with the assumption that our place in the Universe is not a special one and any other observer at any other location in the

Universe will also see the similar isotropic Universe, homogeneity follows. Therefore, we have good reasons to assume the isotropy and homogeneity in the Universe at sufficiently large scales, also popularly known as *Cosmological Principle*.

The models of the Universe based on the cosmological principle form the simplest solution to the Einstein's field equation of General Relativity, which relates the Einstein tensor (geometrical object) to the Energy-Momentum tensor, and lead to the Friedmann-Lemaitre-Robertson-Walker (FLRW) metric,

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\left[\mathrm{d}\chi^2 + f_K^2(\chi)\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2\right)\right] \tag{1.1}$$

where,

- $a(t)$ is the *cosmic scale factor* which increases with the *cosmic time* (t) and normalised to unity today ($a(t_0) = 1$). Hence, the cosmic scale factor (or scale factor henceforth) parametrizes the size of the Universe at cosmic time t relative to today. The cosmic time, and the scale factor characterise an epoch in the cosmic history.
- c is the speed of light.
- χ is the radial *comoving* coordinate.
- θ, ϕ are the angular coordinate on a unit sphere.
- $f_K(\chi)$ relates the distance to the circumference that depends on the curvature parameter K as,

$$\begin{aligned} f_K(\chi) &= K^{-1/2}\sin(K^{1/2}\chi) \quad (K > 0) \\ &= \chi \quad (K = 0) \\ &= (-K)^{-1/2}\sinh((-K)^{1/2}\chi) \quad (K < 0). \end{aligned} \tag{1.2}$$

This metric is written in co-moving coordinate such that the coordinates of the observer are expanding with the Universe and therefore, it does not feel any acceleration.

1.1.2. Cosmological redshift

Consider a source at time t in the cosmic history emitting light which is reaching us today. Due to the expansion of the Universe, the wavelength of the light observed ($\lambda_{\rm obs}$) will be different (stretched or redshifted) from the wavelength at which it was emitted ($\lambda_{\rm emit}$) by a factor determined by the rate Universe expanded during the travel time of the photons, i.e.,

$$(1+z) \equiv \frac{\lambda_{\rm obs}}{\lambda_{\rm emit}} = \frac{a_{\rm obs}}{a_{\rm emit}} \tag{1.3}$$

and as per convention, we define $a_{\rm obs} = a_{\rm t_0} = 1$, we have,

$$(1+z) = \frac{1}{a(t)} \quad \Rightarrow \quad \mathrm{d}z = \frac{\mathrm{d}a}{a^2} \tag{1.4}$$

where z is known as the *cosmological redshift* (or hereafter redshift) of the source which also describes an epoch in the cosmic history of the Universe similar to the scale factor a and the cosmic time t. In many texts, the terms scale factor a, redshift z and cosmic time t are used interchangeably to characterize an epoch in the cosmic history of the Universe. The redshift of a source is a direct observable and has a wide variety of applications in cosmology.

1.1.3. Kinematics

Consider an expanding sphere. A small element on this sphere is marked by its position $\mathbf{x}$. Because the expansion is expected to be radial, the position of the element at a later time t is $\mathbf{r}(t) = a(t)\mathbf{x}$. Differentiating this equation with respect to the time, we get,

$$\mathbf{v}(r,t) = \dot{\mathbf{r}} = \dot{a}\mathbf{x} = \frac{\dot{a}}{a}\mathbf{r}(t) \tag{1.5}$$

We define the *Hubble parameter* or the *expansion rate* as,

$$H(t) \equiv \frac{\dot{a}}{a}; \quad H_0 \equiv H(t_0) = \dot{a}(t_0) \tag{1.6}$$

Hubble parameter today, i.e. H_0, is referred to as *Hubble constant*.

1.1.4. Cosmic time and distances

It is now straightforward to calculate the age of the Universe at time t (scale factor $a(t)$) after the big bang,

$$H(a) = \frac{\dot{a}}{a} = \frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}t}, \tag{1.7}$$

which gives,

$$t(a) = \int_0^a \frac{\mathrm{d}a'}{a'H(a')}. \tag{1.8}$$

This equation gives the age of the Universe when its size was scaled to the scale factor a with respect to the present. Hence, both t and a define an epoch in the cosmic history of the Universe. At $a = 1$ (i.e., today), we define the age of the Universe t_0.

As light rays follow null geodesics ($\mathrm{d}s^2 = 0$), using equation 1.1 we have,

$$c^2\mathrm{d}t^2 = a^2\mathrm{d}\chi^2 \Rightarrow c\mathrm{d}t = -a(t)\mathrm{d}\chi \tag{1.9}$$

where the negative sign indicates that we are looking in our backward light cone.

We can calculate the comoving distance to the source whose photons reach us today by integrating equation 1.9,

$$-\mathrm{d}\chi(a) = \frac{c\mathrm{d}t}{a} = \frac{c\mathrm{d}a}{a^2 H(a)} \quad \Rightarrow \quad \chi(a) = \int_1^a \frac{\mathrm{d}a}{a^2 H(a)} \tag{1.10}$$

Thus, one can use t, a, z and χ as variables to characterize the radial distance of a source that we see today or the epoch at which the light was emitted by the source.

Suppose one knows the size (d) of a source at redshift z, then she can measure its distance by measuring the angle subtending to the observer (θ), this distance is referred to as *Angular Diameter Distance* and is defined as,

$$D_{\mathrm{ang}}(z) = \frac{d}{\theta} = \chi a = \frac{\chi}{1+z} \tag{1.11}$$

the multiplicative factor a is due to the fact that the Universe was smaller by a factor of a when the light was emitted. Angular diameter distance is the distance to the source that we observe today if the Universe had stopped expanding when the photons were emitted from the source.

Similarly, if one measure the flux (S) from a source and model its luminosity (L), it is possible to infer the distance to the source, referred as *Luminosity Distance*, and given by,

$$D_{\mathrm{lum}}(z) = \sqrt{\frac{L}{4\pi S}} = \frac{\chi}{a} = (1+z)\chi \tag{1.12}$$

The luminosity scales as $1/a^2$ where one a is the contribution due to the redshift of the photons and the other comes from time dilation. Hence, the luminosity distance scales as $1/a$.

The era of modern cosmology began with an observational breakthrough by Edwin Hubble in 1928 stating that the galaxies move away from us with a radial velocity ($\mathbf{v}$) which is proportional to their distance from us($\mathbf{r}$) [28]. The proportionality constant is known as the Hubble constant (H_0) i.e., $\mathbf{v} = H_0\mathbf{r}$ is also known as Hubble's law. The measurement consists of the redshift of the galaxies, that Hubble referred to as *nebulae*, and their luminosity distance. For local Universe, it follows a linear relation. This breakthrough along with the homogeneity accounts as the observational evidence for the expansion of the Universe and puts an end to the static world models. The Hubble constant H_0 usually parametrized as, $H_0 = h\ 100\ \mathrm{km\ s^{-1}\ Mpc^{-1}} \approx h/10^{10}\ \mathrm{year}^{-1}$, where, h is a dimensionless number of the order unity which gives the Hubble constant in units of 10 billion years inverse.

To measure the luminosity distance to a source, we need to know its intrinsic luminosity. One class of such objects are the Supernovae-Ia, these are also known as the *standard candles* in the Universe [9, 16]. They don't really have same intrinsic luminosity, as the name can be misleading, but empirically the relation between their peak luminosity and width of the light-curves has a very tight correlation and therefore it is possible to infer the peak-luminosity of such objects by looking at their light-curves. Two teams, lead by Saul Perlmutter [56] and Brian P. Schmidt [67],

independently measured these distances along with their redshift and concluded that the observed fluxes of the high redshift Supernovae-Ia require expansion of Universe at an accelerated rate. In 2011, Saul Perlmutter, Brian P. Schmidt and Adam G. Riess shared a Nobel prize in physics for this remarkable breakthrough.

1.2. The concordance model

Plugging FLRW metric (equation 1.1) into the Einstein's field equation, we get an independent dynamical equation for the scale factor a,

$$H^2(a) \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2}, \tag{1.13}$$

where

- ρ and p are the density and pressure respectively for the matter content of the Universe that must follow a homogeneous and perfect fluid dynamics.
- K is the curvature parameter and follows $K = 0$ for a flat Universe.
- G and c are the Gravitation constant and speed of light respectively.

Equation 1.13 is also popularly known as the *first Friedmann equation*.

ρ may contain more than one component which change with time as the Universe expands. One can distinguish three main components: matter, radiations and vacuum energy. In order to determine the evolution of the densities of these components, let's start with the first law of thermodynamics in an adiabatically expanding Universe, i.e., $\mathrm{d}U = -p\mathrm{d}V$, which can be written in comoving framework using scale factor as,

$$\mathrm{d}(\rho c^2 a^3) = -p\mathrm{d}(a^3). \tag{1.14}$$

One can also relate ρ and p with an *equation of state*:

$$w = \frac{p}{\rho c^2} \tag{1.15}$$

w is known as the equation of state variable for the respective component of the energy density of Universe. For pressure-less matter, $w_m = 0$, for radiation $w_r = 1/3$ where for vacuum energy $w_v = -1$. Using these values in equation 1.15 and combining with equation 1.14, we derive the evolution of densities of various components as,

$$\begin{aligned} \rho_m(t) &= \rho_{m0} a^{-3} \\ \rho_r(t) &= \rho_{m0} a^{-4} \\ \rho_\Lambda(t) &= \rho_{\Lambda 0} \end{aligned} \tag{1.16}$$

where, ρ_{m0}, ρ_{r0} and $\rho_{\Lambda 0}$ are the respective densities today. These results are very intuitive as well. The matter density decreases with the volume of the Universe (a^3)

as the total matter in the Universe remains constant. Similar behaviour applies for radiation plus an additional factor of a due to the redshift of the photons, hence radiation density decreases with a factor of a^4. Finally, the energy density of the vacuum (Λ) stays constant, as more and more vacuum is created with the expansion of the Universe. Therefore, we can write the evolution of the total density of the Universe as,

$$\rho(t) = \rho_m(t) + \rho_r(t) + \rho_\Lambda(t) = \frac{\rho_{m0}}{a^3} + \frac{\rho_{r0}}{a^4} + \rho_\Lambda \tag{1.17}$$

The curvature parameter K in equation 1.13 changes sign with the expansion rate, $K = 0$ is a limiting case. Employing equation 1.13 today, with $K = 0$ yields,

$$\rho_{cr} \equiv \rho_0 = \frac{3H_0^2}{8\pi G} \tag{1.18}$$

ρ_{cr} is defined as the critical density of the Universe. It characterises the total density that is needed to keep the Universe spatially flat. One can normalise the densities of various components of the Universe with critical density to define dimensionless density parameters of the order unity,

$$\Omega_m \equiv \frac{\rho_{m0}}{\rho_{cr}}; \quad \Omega_r \equiv \frac{\rho_r}{\rho_{cr}}; \quad \Omega_\Lambda \equiv \frac{\rho_\Lambda}{\rho_{cr}}; \tag{1.19}$$

It also follows that for a spatially flat Universe, total normalised density parameter Ω_0 is unity,

$$\Omega_0 = \Omega_m + \Omega_r + \Omega_\Lambda = 1 \tag{1.20}$$

Any deviation of this parameter from unity is the indication of spatial curvature of the Universe. Finally, we can re-write the first Friedmann equation in terms of these parameters as,

$$E^2(a) \equiv \frac{H^2(a)}{H_0^2} = \left[\frac{\Omega_r}{a^4} + \frac{\Omega_m}{a^3} + \Omega_\Lambda - \frac{Kc^2}{a^2 H_0^2}\right] \tag{1.21}$$

This equation is extremely important in cosmology, it gives the evolution of the Hubble parameter. It is also used to compute cosmic time, distances, volumes etc. at an epoch based on the scale factor or redshift for given cosmological parameters. This is called the *concordance model*. The current constraints on the cosmological parameters indicates: $\Omega_m \sim 0.3$, $\Omega_\Lambda \sim 0.7$ and $h \sim 0.7$ [60].

1.2.1. Equation of state of dark-energy

Differentiating equation 1.13 and making use of equation 1.14, we get the equation of motion of the Universe, also referred as the second Friedman equation,

$$\left(\frac{\ddot{a}}{a}\right) = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right), \tag{1.22}$$

Employing this equation, it is clear that in order to have an accelerated expansion of

the Universe $\ddot{a} > 0$, there must be a component of the energy density of the Universe that follows,

$$p < -\frac{\rho c^2}{3} \tag{1.23}$$

Such a component resembles the so-called *Dark-Energy*. It exhibits negative pressure large enough to support the accelerated expansion of the Universe. The cosmological constant Λ or the vacuum energy is a special case of the dark-energy where $p = -\rho$. If we describe in terms of equation of state (equation 1.15), it indicates that for dark energy, $w_{DE} < -1/3$, whereas for cosmological constant it's -1. In the rest of the text and also in the subsequent chapters, w indicates the equation of state of dark-energy, unless stated otherwise.

It is also not necessary to assume that equation of state of dark-energy is constant over time. Many different parametrizations were proposed for an evolving equation of state with cosmic time $w(a)$. A popular and widely used parametrization is the so-called CPL (Chavallier and Polarski 2001 [13]; Linder 2003 [40]) parametrization,

$$w(a) = w_0 + (1 - a)w_a \tag{1.24}$$

where, w_0 is the equation of state today and w_a is the derivative with the scale factor.

1.2.2. Thermal History of the Universe

Our current understanding of the Universe from various cosmological probes indicates that $h \sim 0.7$ that gives the age of the Universe $\sim$13.7 Gyr, $\Omega_m \sim 0.3$ which is dominated by the cold dark-matter that interacts only gravitationally, and $\Omega_\Lambda \sim 0.7$, with $w = -1$, which is the dominant component of the energy content of the Universe and applies negative pressure, and hence is responsible for the accelerated expansion of the Universe. Finally, $\Omega_r \sim 10^{-4}$ and is mainly dominated by a background radiation at temperature ~ 2.7 K that can be observed at microwave wavelength in all directions in the sky. This radiation is known as the *Cosmic Microwave Background* (CMB) radiations and shows a nearly perfect black-body spectrum.

Using equation 1.17 it can be seen that in the past matter were dominated over the vacuum energy and if we go long back in the past, radiations were the dominant component. Also radiation loses energy due to the redshift, the CMB photons in the past must be very energetic and therefore hot (as the expansion of the Universe preserves the black-body spectrum). Therefore the temperature of the CMB photons (and hence that of the Universe) drops linearly with the scale factor or the redshift,

$$T_{\rm CMB}(z) = T_0(1 + z) \tag{1.25}$$

At very early times in the history of the cosmic expansion, the scale factor was very small, close to zero, and hence the Universe must have been very hot. Just after the Big Bang, the temperature was so high that the kinetic energy of protons would be large enough to overcome the Coulomb barrier and fuse with other protons, whereas the neutrons had no such barrier. So, it is expected that

light nuclei were synthesised during the early times in the Universe, this process is also known as *Big-Bang Neucleosynthesis* (BBN). As the Universe expanded, it cooled down rapidly and thus the process of BBN could have only happened up to certain time scale. During the first three minutes of the cosmic evolution, all the light mass nuclei were formed and then the temperature dropped below the binding energies of the lightest nuclei and thus no further BBN could happen. During this time, ^{4}He were formed and its mass fraction reached to about 25 percent and only a small fraction Deuterium, Lithium and Beryllium were formed. The BBN is a very important aspect of standard model of cosmology, however, we are not discussing it here in more details. For a thorough review on BBN see [18, 86]. The abundance of these light elements can be calculated with a nuclear reaction network and can be well matched with the current observations putting strong constraints on the physical baryon density in the Universe. This successful comparison is one of the major achievements of the ΛCDM model.

As the time passes, the Universe cooled down further and the temperature dropped enough ($\sim$ 3000 K) such that electrons and protons could combine together to form neutral hydrogen atoms. This process is called *Recombination*. During the recombination, the CMB photons scattered for the last time, sometime it is also referred as the *last scattering surface*. After recombination photons get decoupled from the baryon-photon fluid and stream freely which can be observed today as the CMB. This happen ∼380,000 years after the big bang or at redshift ∼1100. It was not an instantaneous process, the Universe became neutral in a short interval of time. When we look at the CMB sky, we actually see this ionization front. In 1965, A. Penzias and R. Wilson discovered the CMB accidentally when they noted an excess antenna temperature of about 3K that they could not remove from any known source [18]. They were given a Nobel prize in physics in 1978. In 1992, Cosmic Background Explorer (COBE) satellite which scanned the full CMB sky, found fluctuations in CMB temperature of the order 10^{-5} [19, 75]. These anisotropies originate from the matter density fluctuations at the time of the recombination. COBE also observed a perfect black-body spectrum in CMB. In 2006, when another CMB all sky survey Wilkinson Microwave Anisotropy Probe (WMAP) [32] confirmed the findings of COBE with much higher signal to noise ratio, COBE team leaders George Smoot and John Mather were awarded with a Nobel Prize in physics. Very recently these anisotropies were studied with much better signal to noise ratio and up to much smaller scales by Planck satellite, putting tightest constraints on the cosmological parameters.

After recombination, the Universe became neutral. However, current observations show that the Universe is fully ionized up to redshift $\sim$ 6. Therefore, sometime between redshift $\sim$ 1100 and $\sim$ 6, the Universe must have been re-ionized either by the first stars or the quasars. As it is not an instantaneous phenomenon, the re-ionization of the Universe must have started much earlier redshift 6. The most recent results from CMB observations indicates that the redshift of the re-ionization is $\sim$ 8 [60]. The period between the recombination (redshift $\sim$ 1100) and the re-ionization (redshift $\sim$ 8) is called the *Dark Ages*.

1.2.3. Cosmic inflation

Despite the success of the ΛCDM model of cosmology, there are few caveats. Two major conceptual problems in the framework comprise its main drawbacks:

- Flatness problem: It is also referred as *fine tuning* problem. It states that in order to have the Universe spatially flat today, i.e. the curvature density close to zero (maybe within a margin of few percent), the curvature parameter in the early Universe (say $z \sim 10^{10}$) must be extremely close to zero, of the order 10^{-15}. If this condition is not satisfied, the Universe would have re-collapsed long ago that there would have been no time for the stars or planetary systems to form or it would have expanded significantly faster than the Universe we live in such that it would have prevented the formation of the structures in the Universe. In both cases, it wouldn't be possible for life of any form to exist to study the flatness of the Universe. Hence an extreme fine tuning of the density parameters is needed at early times for life to exist today. For a review see [33].

- Horizon problem: Due to the finite speed of light, we can only observe a finite part of the Universe, only those regions from where the light can reach us in time t_0. So, there is a part of the Universe beyond which we cannot see, this boundary mark the size of the *observable Universe* and known as the *horizon*. This also implies that the horizon size must have been smaller in the past. If one calculates the size of the horizon at the time of recombination within ΛCDM framework and divides it by the angular diameter distance to the CMB, the angular size of the horizon at the time of the recombination comes out to be $\sim 1^\circ$. This means that the regions in the sky (or CMB sky) which are separated by more than a degree were never causally connected. Still as we observe in the CMB temperature fluctuations, the temperature of whole CMB sky is almost the same and relative difference is of the order 10^{-5}. How come the regions in the sky are so isotropic when they never exchanged information? For a review see [33].

One can always assume the initial condition to be such that the above two conditions follow. But this does not explain anything and is highly unlikely. In 1980's a new model was developed known as the *inflationary model* (or inflation; for a review see [6, 10, 84]). It presumes that at a very early time the vacuum energy was much higher than today and dominated the Hubble expansion, resulting in an exponential expansion of the Universe. This inflationary phase ended when the vacuum energy transformed into matter and radiation via a process referred to as *reheating [6].*

The theory of cosmic inflation solves the flatness and horizon problem (and few others as well). Due to the very fast exponential expansion in early times, the regions which were not causally connected got connected. This immediately solves the horizon problem. Also due to such rapid expansion, any initial curvature which was not fine tuned could have been straightened out and it got highly flat. This solves the flatness problem.

The scenario of the inflation, the exponential expansion of the Universe at an early time, solves many problems in ΛCDM, yet the physical mechanism is unknown. A very big problem is that how and when the inflationary phase came to an end and why not at any other time. Nevertheless, inflation is one of the important and plausible scenario in the framework of ΛCDM model. Because inflation follows spatial flatness of the Universe, we can drop the K terms in the Friedmann equations in the rest of the text. For most recent constraints on the inflationary models, see [58].

1.3. Structure Formation in the Universe

An absolutely homogeneous Universe would be very easy to describe mathematically, but it would be devoid of any interesting physical phenomenon that we observe around us. Thanks to the inhomogeneities in the Universe, we exist! The Universe shows these inhomogeneities or structures on various scales up to $\sim 200\ h^{-1}$Mpc. The seeds of these inhomogeneities were the tiny quantum fluctuations [70] , which as per our concordance model, are amplified up by the *inflation* and are transformed into small density perturbations which can be observed in the CMB. During the early dark ages, these perturbations were still very small and as a result of gravitational instability, these small perturbations grew and formed the large-scale structures of the Universe. So, the history of the structure formation in the Universe, as described in the concordance model, can be studied in two parts: (i) *Linear theory*, when the size of the perturbation were small and higher order perturbation terms can be ignored and (ii) *Non-linear theory*, when the size of the perturbations grows significantly larger that higher order terms cannot be ignored.

1.3.1. Linear theory

The earliest perturbations we see in nature are the fluctuations in the CMB due to the perturbations in the matter density field. Because these perturbations are tiny ($\sim 10^{-5}$), the higher order perturbation terms can be neglected and hence, linear theory can describe growth of structures at very high redshift (early dark ages). Secondly, the local Universe is homogeneous at scales larger than $\sim 200\ h^{-1}$Mpc and therefore, if we smoothen the distribution of matter at scales larger than this, it looks homogeneous and hence the perturbations are small. Therefore, linear theory can also be applied at large scales. So, linear theory of the structure formation and evolution can be applied to the large scales at all redshift and to small scales only during early stages of the dark ages. Within these regimes, the theory of structure formation can be described analytically.

We start by defining the density contrast parameter which is the relative deviation of the density from the mean background density of the Universe,

$$\delta(r,t) = \frac{\rho(r,t) - \bar{\rho}(t)}{\bar{\rho}(t)}, \tag{1.26}$$

where, $\bar{\rho}(t)$ is the mean matter background density of the Universe at that epoch.

Assuming a matter dominated Universe (or the pressure due to radiations is zero), which is a good approximation at the high redshift, with dark-matter as the dominant component being collision-less. So, we can approximate the matter as a pressure-less fluid which is fairly valid at large scales. The fluid equations for vanishing pressure are: (i) *Continuity equation*, describing the conservation of matter; (ii) *Euler equation*, describing the equation of motion for the fluid and finally (iii) *Poisson equation* describing the gravitational field. Solving the system of these equations in linear regime (i.e., neglecting all higher powers of δ) in the comoving coordinates will give the so-called *Growth Equation*:

$$\frac{\partial^2\delta}{\partial t^2} + \frac{2\dot{a}}{a}\frac{\partial\delta}{\partial t} - \frac{3H_0^2\Omega_m}{2a^3}\delta = 0 \tag{1.27}$$

This equation is a reasonable approximation at large scales, where the perturbations are small, it and completely describes the distribution and evolution of perturbations δ. It is valid for both real and Fourier space $\tilde{\delta}$ where we use the convention,

$$\tilde{\delta}(k,t) = \int_{-\infty}^{\infty} \delta(r,t)\exp(-2\pi\imath kr)\mathrm{d}r. \tag{1.28}$$

In equation 1.27, δ is a function of both space and time, but the derivative is only with respect to the time and there is no space term in the coefficients, it's an ordinary second order differential equation. The solution to this equation can be separated in spatial and temporal components, where the temporal part gives the growth of the structures, $D_+(t)$, and therefore,

$$\frac{\partial^2 D_+(t)}{\partial t^2} + 2H(a)\frac{\partial D_+(t)}{\partial t} - \frac{3H_0^2\Omega_m}{2a^3}D_+(t) = 0 \tag{1.29}$$

$D_+(t)$ is also known as the *Growth Factor*. It describes the evolution of the linear density fluctuations with cosmic time. Once the initial conditions are defined, this equation can be used to calculate the growth of linear perturbations.

If we also include the contribution from the fluid pressure etc., these equations remains the same except for the last term,

$$\ddot{\tilde{\delta}} + 2H(a)\dot{\tilde{\delta}} = \left(\frac{3H_0^2\Omega_m}{2a^3} - \frac{c_s^2k^2}{a^2}\right) \tag{1.30}$$

where, the second term on the right is the pressure term with c_s as the speed of sound. This term is only important for large values of the wave-vector (k) and hence for the small scales. Therefore, at large scales pressure is not important, and hence gravity will dominate and vice-versa. This defines a characteristic length scale below which pressure becomes important, *Jeans length* $\lambda_j \equiv c_s\sqrt{\pi/G\rho}$. As the Universe is expanding, the density of the Universe and also the speed of sound is changing and so does the Jeans length.

Now that we have described the evolution of density field with cosmic time, with and without pressure term, one can predict the fluctuations at any time given the

initial conditions. Describing the initial conditions for the density fluctuations immediately raises the question: how to characterize a function describing the density field δ in space at any time in cosmic history? One cannot expect any physical theory to do that, as the density field in initial conditions is a random field, at best one can hope is to describe its statistical properties.

The simplest of the statistics is the two-point function. Suppose there is a completely random distribution of galaxies in the Universe without any deterministic force. Now, given a galaxy at origin, what is the probability of finding a galaxy at a distances r_1 and r_2; it must be the same. However, due to the Gravitational force, galaxies attracts each other, and hence there is this an excess probability to find a galaxy closer to another galaxy. So to say, the probability of finding another galaxy at r_1 is larger than at r_2 if $r_1 < r_2$.

This excess probability can be modelled as the *two-point correlation function* (2PCF) $C(x, y)$ which is defined as,

$$\langle \delta(x)\delta^{\star}(y)\rangle = C(|x-y|) \tag{1.31}$$

The 2PCF is a function of the distance separation between two points in the space. It is more convenient to work in Fourier space and hence we define the Fourier transform of the 2PCF as the power spectrum ($P(k)$) of the density field or commonly known as *matter power spectrum*,

$$P(k) = \langle |\tilde{\delta}(k)|^2 \rangle \tag{1.32}$$

where, k is the wave-vector at the corresponding scale. If we define the matter power spectrum of the initial density field and combine it with the growth factor, we have a complete description of the evolution of linear density field in the Universe.

To define the initial power spectrum we have two important facts to take into account:

- First: at very early times, all relevant length scales were larger than the size of the horizon, there were no characteristic scales in the units of which k could be measured. The only mathematical function for length that does not require a characteristic length scale is the power law. Therefore we expect the initial power spectrum to be a power law with some index (say n_s), i.e., $P(k) \propto k^{n_s}$. n_s is also known as the *spectral index*.

- Second: the growth of the density fluctuations depends on their length-scale as compared to the horizon size at that epoch. Also the growth of various perturbations would be different had they entered the horizon in a radiation-dominated phase or a matter dominated phase etc.. Therefore, a transfer function $T(k)$ is defined that accounts for the fact that the growth of small-scale perturbations is suppressed relative to those which enter the horizon only after matter domination. Therefore, we have $P(k) \propto T^2(k)$

Thus we can define an initial power spectrum, which is also linear in nature as,

$$P_{\rm Lin}(k) = Ak^{n_s}T^2(k) \tag{1.33}$$

where, A is the normalisation factor. Therefore, we can re-write the linear power spectrum at any epoch or redshift z as,

$$P_{\rm Lin}(k,z) = P_{\rm Lin}(k) D_+^2(z) \tag{1.34}$$

The shape of the linear power spectrum is determined by the parameter n_s and transfer function. There is a different form of normalisation in practice such that if one counts galaxies in a sphere of radius 8 h^{-1}Mpc, then the average relative error is close to unity, i.e.,

$$\langle \delta^2(8h^{-1}{\rm Mpc}) \rangle \approx 1 \tag{1.35}$$

So, we define the variance of the smoothened density field as,

$$\sigma^2(R) = \int \frac{d^3k}{(2\pi)^3} |\tilde{W}(k,R)|^2 P(k) \tag{1.36}$$

where, $\tilde{W}(k,R)$ is the Fourier transform of the top-hat function smoothened at scale R and is given by,

$$\tilde{W}(k,R) = 3\frac{\sin(kR) - kR\cos(kR)}{(kR)^3} \tag{1.37}$$

and therefore we have,

$$\sigma^2(8h^{-1}{\rm Mpc}) \equiv \sigma_8^2 \approx 1 \tag{1.38}$$

σ_8 is the parametrisation in practice to normalize the power spectrum. σ_8 along with the spectral index n_s completely describe the power spectrum for the initial density field.

1.3.2. Non-Linear theory

At late times, the approximation $\delta \ll 1$ is not valid at small scales. So, the analytical description for the structure formation and evolution becomes impossible. The solution then would have to rely on higher order perturbation theory (PT), physical approximations to the distribution of the matter in the Universe (e.g., the halo model), semi-analytic approaches (e.g., fitting functions, modified halo model etc.; see [74, 80]), or simulations. The higher order PT are quite useful for many applications but it also starts to fail as δ becomes of order unity (for a comparatively old but thorough review see [17]). The physical approximation methods turn out to be very powerful in predicting the observables approximately but since they are based on strong assumptions which are certainly not always true, they either become unreliable when precision is required, or have to be calibrated on simulations. Finally, simulations are the ultimate solution but they are limited by the volume factor and resolution, and are computationally expensive.

1.3.2.1. N-body simulations

With advancement in the computational power in the last decade, it is now possible to simulate the Universe at cosmological volumes to an accuracy up to the scales where only gravity is important in structure formation. These scales are highly non-linear and are very difficult to describe by other methods. The smaller scales, where the baryonic interactions become important, have only limited success, but the progress has a very steep slope. For simplicity, in this section, I shall restrict the discussion about dark-matter only (DMO) simulations, also popularly known as $N-$body simulations. Baryonic processes in structure formation will be discussed in section 1.3.4.

An N-body simulation characterizing a Universe consists of only dark-matter (and no baryons) which interacts only gravitationally. Such a simulation is useful in studying the structure formation at scales large enough that baryonic processes are not important and small enough that the clustering processes are non-linear and difficult to describe analytically. The basic principles of such simulations are described below.

It is not possible (or at least very expensive) to simulate a volume equivalent to the observable Universe at high enough resolution to study the evolution of individual halos. Assuming that the Universe is homogeneous at cosmological scales (or the scales of the largest structures), the simulation volume must at least include these structures. So a cube with side length $\sim 200h^{-1}$Mpc or larger is needed to simulate a true representative of the Universe. Further, the particles close to the boundaries of the cube must also feel the gravitational pull from outside the box and hence one cannot consider the space outside the box to be empty. Since the Universe is homogeneous at scale larger than the size of the box defined, one can assume the *periodic boundary conditions*. For example if a particle leaves the box from left side, it re-enter the box from the right side. Also, the particles close to the left boundary will feel the gravitational pull towards left similar to as if the distribution of the particle towards the left is similar to the right side boundary.

One very important ingredient of a simulation is the initial conditions which are set to very high redshift. Let the particles evolve with gravity up to redshift zero. The best way to set the initial condition is to put all the particles in a uniform 3D grid and then displace each particle with a displacement field to match the initial power spectrum (linear).

Another important aspect of a simulation is its resolution. This depends on the number of particles or the mass of each dark-matter particle (at least one of them is free parameter to be fixed at the beginning). The smaller the mass of the individual dark-matter particle, the larger will be the total number of particles in order to match the critical density of the Universe today and hence, one can resolve smaller scale structures. There is always a trade-off between large volumes and higher resolutions.

Recently, Skillman et al. [73] has made available a very ambitious N-body simulation of box size $8h^{-1}$Mpc with a billion particles. The total size of the output is close to 55 TB. Schneider et. al. [68] studied various cosmological N-body simulation codes (RAMSES,PKDGRAV and GADGET2) and computed power spectrum giving full analysis of volume and resolutions required for a precise measurement of the

matter power spectrum. Heitmann et al. [23–25, 34] has made a available an accurate matter power spectrum emulator based on 38 cosmological N-body simulations.

Analysing the output of an $N-$body simulation is a fun exercise and a lot can be learnt from it. First, merely the visualisation of the evolution of particle positions gives us an insight and a better understanding of the formation of cosmic web and large scale structures of the Universe. Also, one can calculate the non-linear matter power spectrum and its evolution using equation **??** and constrain the theory of structure formation. One very important application is to search for dark-matter halos in the simulations and compare it with other physical theories to put constraints on their parameters and also to compare with observations. However, the relation between dark-matter halos and luminous galaxy distribution needs further considerations like galaxy bias etc. and is beyond the scope of this work.

1.3.2.2. The halo model

The halo model [43, 45, 55, 71] is one of the more successful, analytic and physical frameworks to describe the clustering and growth of structures in the Universe. In this framework, all the matter in the Universe is assumed to be in the form of spherical halos whose radius are defined by a density threshold. Generally R_{200}, which is the distance from the centre of the halo where the density inside the sphere drops to 200 times the mean matter density of the Universe, is used. The distribution of mass inside the halos are assumed to follow a radial density profile, which depends on its mass, truncated at R_{200} or at virial radius ($R_{\rm vir}$).

The assumptions of the halo model don't hold in details but are well approximated and hence the estimators for the matter power spectrum are well in agreement (nearly 20% at non-linear scales) to the more accurate simulations. However, due to these invalid approximations, it is not possible to achieve sub-percent level accuracy with this estimator.

There are particularly four ingredients to model the power spectrum in the halo model framework:

- Linear power spectrum $P_{\rm Lin}(k)$.
- Radial density profile of the halos. Usually for dark-matter only Universe, the so-called Navarro-Frenk-White (NFW; [53]) is used as a Universal profile, which is completely characterized by the mass of the halo,

$$\rho(r) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2} \tag{1.39}$$

  where, r_s and ρ_s are the scale radius and characteristic density at scale radius respectively.
- Halo mass function $f(\nu)$: a functional form defining the number of halos as a function of halo mass that follows:

$$\int_0^\infty f(\nu)\mathrm{d}\nu = 1 \tag{1.40}$$

where, ν is called the peak-width of the density peak of the halo characterizing its mass. This condition implies the assumption of the halo model that all the mass in the Universe is inside the halos.

- Halo bias $b(\nu)$: As the halos are biased tracers of the matter, one important ingredient is the description of the bias. This term is important while evaluating the correlation between different halos. In the halo model framework, halo bias follows,

$$\int_0^\infty f(\nu)b(\nu)\mathrm{d}\nu = 1 \tag{1.41}$$

Once these four ingredients are defined, it is very straightforward to model the power spectrum. In this framework, the total matter power spectrum is split into two main contributions,

$$P(k) = P_{1h}(k) + P_{2h}(k) \tag{1.42}$$

The first term on the right is referred to as the *one-halo* term ($P_{1h}(k)$) and gives the correlations of the dark-matter particles inside the halo and is computed as the auto-convolution of the halo profile which dominates at the smaller scales. The second term is known as the *two-halo* term ($P_{2h}(k)$) and gives the correlation between different halos. As the dark-matter halos are biased tracers of the mass distribution, this term has a contribution from the halo bias and it dominates over the larger scales.

The halo model is discussed in much more details in chapter 3 and 4 but for a more extensive review see [14].

1.3.3. Covariance matrix of matter power spectrum

As the 2PCF and power spectrum are statistical quantities, there is a statistical limit to which it can be measured in cosmological surveys or simulations. This limit is expressed as the covariance matrix, defined as,

$$\mathbb{C}_{k,k'} = \langle P(k)P(k')\rangle - \langle P(k)\rangle\langle P(k')\rangle \tag{1.43}$$

The estimator for the full covariance matrix of the matter power spectrum can be modelled as the contributions from three broad parts:

$$\mathbb{C}(k,k') = \mathbb{C}^G(k,k') + \mathbb{C}^{NG}(k,k') + \mathbb{C}^{SSC}(k,k') \tag{1.44}$$

where, the three terms on the right are the Gaussian, Non-Gaussian and super-sample covariance contributions respectively.

- Gaussian part $\mathbb{C}^G(k,k')$: The matter power spectrum is computed by averaging over the modes in spherical shells in the Fourier space. It is understood that if one averages over more number of modes, the underlying error on the average

will be smaller and vice-versa. So, the Gaussian part of the covariance matrix gives this contribution and has Poissonian structure.

$$\mathbb{C}^G(k,k') = \frac{2}{N(k)}\delta_{k,k'}P^2(k) \tag{1.45}$$

where, $N(k)$ is the number of modes corresponding to the wave-vector k. One can either count the number of modes while averaging over the power spectrum or can compute this as,

$$N(k) = 4\pi k^2 dk \left(\frac{L}{2\pi}\right)^3 \tag{1.46}$$

where, L is the box size and dk is the bin width. The Gaussian part contributes only to the diagonal elements of the covariance matrix and provides no information about the correlated errors between different modes. As we have only one Universe to observe, large scale modes are very few, and hence the covariance of the matter power spectrum is dominated by the Gaussian part for small k. It is also popularly known as *cosmic variance* or *sample variance*.

- Non-Gaussian part, $\mathbb{C}^{NG}(k,k')$: The non-Gaussian covariance term is induced by the non-linear growth of the density perturbations and arises from the connected 4-point function or the trispectrum due to the fact that the phases of $\tilde{\delta}(k)$ are not random. This term gives the correlated errors between different wave-vectors and also contributes to the off-diagonal terms of the covariance matrix. This contribution is vanishing in the early times if the initial density field is Gaussian and so all modes evolve independently. So, when the non-linearity enters and modes become correlated, the dispersion of the power spectrum increases [47].

- Super-sample covariance (SSC), $C^{SSC}(k,k')$: The SSC is the sampling error due to the coupling to the modes larger than the size of the survey or simulation. It was first reported by [22]. Because this contribution appears constant at all the survey scales, it can be viewed as a curvature term to the survey which can further be mimicked by the change in background density. Therefore this term can be modelled completely by the response of the matter power spectrum to the change in background density [36, 77],

$$\mathbb{C}^{SSC}(k,k') = \sigma_b^2 \frac{\partial P(k)}{\partial \delta_b}\frac{\partial P(k')}{\partial \delta_b} \tag{1.47}$$

where, σ_b is the linear variance.

The accurate quantification of all these terms is crucial in order to perform any likelihood analysis based on observables modelled on the matter power spectrum. As we advance in the quality of data, the smaller scales give more constraining power over cosmological parameter space, the contributions from non-Gaussian part and

SSC become very important and ignoring them may overestimate the constraints on the cosmological parameters and mislead the interpretation by adding biases to the cosmological parameters [30, 35, 72, 79, 92]. In order to estimate the full covariance matrix correctly, matter power spectrum need to be estimated from large sample of cosmological simulations [8, 81, 82].

1.3.4. Baryonic contributions

It is relatively easy to model and describe the clustering of matter that interacts with only one force (gravity). But we exist and are not made up of dark-matter, instead by the ordinary matter which interacts via other forces, in addition to gravity and exhibit pressure. Even the Universe with baryons and dark-matter would behave very similar at large scales, but at small scales where the baryonic processes are important, dark-matter only Universe is not a good description. Many recent works demonstrate different scales at which the baryonic contributions are significant in the matter power spectrum. These scales vary between $k \sim 0.3 - 0.6h\text{Mpc}^{-1}$ which transform into a physical scales smaller than $\sim 10 - 15h^{-1}\text{Mpc}$ (see [87, 88]). Future generation surveys like Euclid[1] [5, 63] and LSST[2] [42] are expected to provide data with very high resolution and are expected to probe scales $k \sim 10h^{-1}\text{Mpc}$ or physical scales of about $0.5h^{-1}\text{Mpc}$ where the baryonic contribution are very significant.

To model the structure formation with baryons is rather a messy task. Analytic calculations can be very complicated and not-reliable. However, computer simulations provide a much better tool to study the formation of structures including baryonic physics. One can always put baryonic matter along with dark matter in the simulations and evolve with governing gravity and baryonic processes which are known to be important. Some of the processes are: Star formation, Feedback processes from AGN or Supernovae etc., Radiative transfer, Adiabatic contraction or expansion of dark-matter due to baryons.

There is very little understanding of star formation processes, other processes can be modelled with some approximations. But carrying out these calculations in big cosmological volumes are very expensive and may still be out of the reach of the current computational abilities. Different hydrodynamical simulations have found a remarkable agreement with various observations. For example, quasar absorption line observations of the Ly-α forest [11, 15, 83], properties of high-column density HI absorbers [46, 61, 62], sizes of disc galaxies [1, 20, 21, 44, 89], intergalactic gas [54, 76] etc.. An example of a state of the art hydrodynamical simulation is the EAGLE project [66].

Another way to include the baryonic processes is to perturb the density profile of the halos with baryonic processes and model the power spectrum using the halo model framework [50]. This approach can be calibrated with simulations or observations but still remains analytical with some free parameters to be measured by data directly. For a similar approach see chapter 4

[1] http://www.euclid-ec.org
[2] http://www.lsst.org/lsst/

1.4. Gravitational Lensing

The bending of light due to the gravitational field of the intervening matter is known as the *gravitational lensing* (or lensing hereafter). If the deflection in the light path is strong enough, an observer will observe multiple images of the same source (also referred as *strong gravitational lensing or SL*), if it is not strong enough, only the distortion in the shape of the source (also referred as *weak gravitational lensing or WL*) will be observed. It is a powerful technique to obtain information about the intervening matter, also known as the *gravitational lens* (lens hereafter). The quantitative description of the lensing phenomenon was formulated only after the description of Einstein's General Relativity (GR), however, it was suspected long before that. Based on Newtonian theory of gravitation, in 1784 John Mitchell mentioned to Henry Cavendish the possibility of bending of light through a gravitational field by an angle $2GM/(c^2\xi)$, where M is a spherical mass like a star and ξ is the impact parameter. GR finds that this number is larger by a factor of 2,

$$\frac{4GM}{c^2\xi} = 1.75'' \left(\frac{M}{M_\odot}\right)\left(\frac{\xi}{R_\odot}\right)^{-1} \tag{1.48}$$

A measurement of this kind was made for background stars during a total solar eclipse in 1919 with sufficient accuracy which made it clear that the GR prediction is indeed correct. This was a tremendous success of Einstein's theory.

In 1924, Chwolson considered a simple geometry where the source is exactly behind the lens and concluded that instead of multiple images, the sources will be seen as a ring around the lens. These are referred to as *Einstein Rings*. In 1936, Einstein considered lensing by a star and concluded that the separation between two images would be of the order of milli arc-seconds and there is no possibility to resolve it. However, more optimistic view was shared by Fritz Zwicky, who, in 1937, proposed to look for lensing phenomenon in galaxies that he called *extragalactic nebulae* and that the image separation to be of the order 10 arc-seconds and certainly be resolved by a telescope. If the source which is multiply imaged by a lens is variable, it is also possible to calculate the difference in the light travel time between two images of the same source. In 1960 Refsdal made such a calculation and pointed out that it depends on the mass of the lens and the distance between lens, observer and source; if they are known it can be used to measure the Hubble constant and infer other cosmological parameters [64].

In 1979, for the very first time CCD detectors replaced the traditional photographic plates which provided high resolution images and thus it is now possible to observe the lensing multiple images which has been theorized for a long time. [90] marked the discovery of first multiply imaged system when they discovered a pair of quasars separated by nearly 6 arc seconds having identical colour, redshift and spectra. This system is known as QSO 0957+561. One year later, another lensing system was observed by [91]. In this system the quasar was lensed into three images, *triple quasar* (PG 1115+080). In 1986, a giant luminous arc was discovered in the galaxy cluster named Abell 370. The arc was a highly distorted image of a high redshift galaxy. Finally in 1988, an Einstein ring was observed in MC 1131+0456

by [26]. By now, more than 400 multiple-image lens systems have been discovered where a bright galaxy or a galaxy cluster is the main lens.

1.4.1. Lensing theory

Consider a gravitational lens at a comoving distance χ (redshift z_l) with sky-projected surface mass density $\Sigma(\theta)$, where θ is a 2D vector to the angular coordinate perpendicular to the line of sight. Due to the gravitational field of the lens, different light rays from a background source will travel different distances before reaching the observer. Therefore, the time delay due to different travel distances between two images of the same source can be split into two main contributions and written as,

$$t(\theta) = \frac{\chi}{c}\left(\frac{\chi'}{2(\chi'-\chi)}(\theta-\beta) - 2\nabla_\theta^{-2}\bar{\kappa}(\theta)\right) \tag{1.49}$$

where, χ' is the comoving distance from the observer to the source and β is the angular positions of the source perpendicular to the line of sight. The first term on the right is due to the geometry of the system and the second term is due to the gravitational potential of the lens. As the total time delay is proportional to the comoving distance to the lens, it is inversely proportional to the Hubble constant. Also, $\bar{\kappa}(\theta) = \kappa\ (\chi'/(\chi'-\chi_l))$ and κ is the surface mass density of the lens defined as,

$$\kappa(\theta) = \frac{4\pi G}{c^2}\frac{\chi}{1+z_l}\left(\frac{\chi'-\chi}{\chi'}\right)\Sigma(\theta) \tag{1.50}$$

$t(\theta)$ is a continuous function and can be visualised as contours of constant time delays in the lens plane. For a source at infinity, $\bar{\kappa} = \kappa$. Images cannot form everywhere, because not every light ray reaches the observer, they will form only at coordinates where $t(\theta)$ has a minima, maxima or a saddle point. These extrema positions can be calculated by finding the roots of the equation,

$$\nabla t = 0 \quad \Rightarrow \quad \beta = \theta - \alpha(\kappa, \theta) \tag{1.51}$$

This equation, also known as the *lens-equation*, relates the true position of the source to the observed position of its images with a deflection angle α that depends on the surface mass density of the lens and position. If this equation has more than one solution, there will be multiple images of the source.

The second derivative of equation 1.49 gives the *deformation matrix* (or the *inverse magnification matrix*),

$$\nabla^2 t = \begin{pmatrix} 1-\kappa+\gamma_1 & \gamma_2 \\ \gamma_2 & 1-\kappa-\gamma_1 \end{pmatrix} \tag{1.52}$$

where, γ_1 and γ_2 are the two components of the shear (or the tidal gravitational force). If this matrix is singular, it gives the curves of the infinite magnifications, also known as *caustics*. For an extensive review on the theory of gravitational lensing see [48, 57, 69].

1.4.2. Applications

Now a days, lensing is a very active field of research which finds application in variety of problems in cosmology and astrophysics. There are various observables due to the lensing phenomenon. For example multiple-images of the source, time delays amongst the multiple-images, shape distortion of the source, change in brightness of the source etc. Each of these observables can be utilised to infer or model various aspects: from the mass distribution of the lens to cosmology. Some of the main applications of lensing are as follows,

- Mass distribution of lens: the deflection of light when passing through the gravitational field of the lens depends on the total mass or the mass distribution of the lens and hence it is independent of the nature of the matter of the lens or its state. Therefore, the deflection is equally sensitive to the dark-matter, non-luminous baryons or the luminous galaxies. It doesn't matter at all if the lens is a virialised structure or a recent merger. This gives an opportunity to study the total mass distribution of the lens, no matter how complex the distribution is. Hence it is an ideal tool for measuring the total mass of the lens, whether dark and luminous, or the exact 2D mass distribution.

- Estimating cosmological parameters: Employing Refsdal's idea, the Hubble constant can be constrained using the time delay measurements between two images of the same source. Weak lensing (WL) by large scale structures can also be used in order to put strong constraints on cosmological parameters. According to the Dark Energy Task Force (DETF; [4]), weak gravitational lensing is the most powerful tool in order to put constraints on the cosmological parameter given the controlled systematics of the WL observables. This will be discussed in more details in section 1.4.4. Furthermore, WL has found application in constraining bias parameters which describe the statistical studies of the distribution of galaxies and dark-matter.

- Natural telescopes: since the gravitational lenses can magnify the sources, it can be used as a natural telescope to study high redshift galaxies and the Universe in general which is otherwise does not lend itself to observation. One of the example of a very ambitious project is the Hubble Frontier Fields (HFF; [41]).

There are a lot of other applications , e.g. - constraining the number density of dense objects, searches for exo-planets etc., these are not discussed in details in this manuscript.

1.4.3. Strong Gravitational Lensing

Strong gravitational lensing (or SL henceforth), refers to the lensing phenomenon where the strength of the gravitational field of the lens is sufficiently strong and the alignment between the source, the lens, and the observer is optimal enough to produce multiple images of the source. From equation **??**, SL regime marked by $\kappa > 1$ and relatively small values of θ. So, the multiple images of a background sources contains information about the projected mass distribution of the lens.

Given the mass distribution of a lens (κ) and position of the source (β), it is straightforward to find out the position of the multiple images (θ), this process is called *forward modelling*. But given a set of multiple images, which are the main observables in SL, it is a highly degenerate process to reconstruct the mass distribution of the lens. The procedure is referred as *lens inversion* or *lens modelling* and mainly consist of reconstructing κ and β (as in equation 1.49).

There are several methods for lens inversion. Many of those procedures put strong constraints on the mass distribution of the lens and assume a functional form and hence are called *parametric* methods. These methods mainly involve to find the optimal values of the free parameters of the functional form of the mass distribution of the lens. For example LENSTOOL[3] [31].

There are some methods which invert the lens with minimal assumptions and without assuming any prior form of the mass distribution of the lens, and are referred as *non-parametric* methods. In these methods, the number of free parameters are often too large, usually some building block of the total mass maps, and hence larger statistical consideration is needed. For example, Pixelens[4] [65], GRALE[5] [37].

Where as parametric models are more efficient, non-parametric models are more accurate and unbiased. In this work we used a non-parametric lens inversion library called as GRALE [37–39] to trace the mass distribution of some very massive galaxy clusters. The building blocks of the mass map are the Plummer spheres, where other choices such as squares and Gaussian spheres are also available in the library, and the total mass map is the super position of these building blocks. GRALE uses a genetic algorithm (for a review see [12]) to find optimal solution to the weights of the building blocks and the resolution of the mass maps is adaptively increased (or decreased) For more detailed documentation see [39].

Another observable in SL is the time delays between multiple images of the source. The light rays from the same source travel different directions and due to the curvature of space time, depending upon the mass distribution of the lens, they travel different distances before they reach the same observer. Because the distances in cosmology depends on the Hubble constant or the expansion rate of the Universe, one can calculate the value of the Hubble constant given measured time delays and mass distribution of the source [64]. However, it is also possible to have the Hubble constant given and put additional constraints on the mass distribution of the lens using measured time delays [51].

1.4.4. Weak Gravitational Lensing

Weak gravitational lensing (or WL henceforth) refers to the phenomenon when the lensing is not strong enough to produce multiple images but strong enough to distort the shape of the source. In WL, deformation matrix is close to unitary matrix. The deformation of the shape of the observed galaxies due to the intervening matter is referred to as *cosmic shear*. This signal is very small, nearly 1-2% of the intrinsic ellipticity of the source and can only be measured statistically under the assumption

[3] `http://projets.lam.fr/projects/lenstool/wiki`

[4] `http://www.physik.uzh.ch/~psaha/lens/pixelens.php`

[5] `http://research.edm.uhasselt.be/~jori/page/index.php?n=Physics.Grale`

that the intrinsic ellipticity of the background galaxies do not have a preferred direction. If one measures the cosmic shear of all background sources behind a lens or mass concentration, it tends to align tangentially towards the centre of the mass concentration.

As the signal of the weak lensing cosmic shear is very small compared to the intrinsic ellipticity of the galaxies, this study did not come into play until the recent observational and technical advancements. Soon after the detection of the giant luminous arc in Abell 370, it was observed that few more objects that are not as stretched as the giant arc but still show high axis-ratio and are aligned tangentially towards the centre of the cluster. They termed it as *arclets* and it was clear that the alignment is due to the gravitational field of the cluster. It was also expected that there are only a few very strong distortion of the shapes could happen like the giant arc, many more smaller distortions can be observed in the background galaxies. [85] reported the first statistical detection of the WL cosmic shear in two lensing clusters. The theoretical framework was further formulated by [29] and it evolved as an active field to study the mass distribution of lensing clusters in the outskirts.

In addition to just studying the local mass concentrations, like galaxies or galaxy clusters, WL can also be used in order to study the statistical distribution of matter in an inhomogeneous Universe. Light rays coming from all redshift are continuously distorted from the matter on its path (figure 1.1). For example, light from high redshift sources are deflected many times as they might witness more mass concentrations in its path, as compared to low redshift sources. If the cosmic shear can be measured for a large ensemble of sources spread over redshift, the statistical properties of these shears can be used in order to study the statistical properties of the cosmological matter distribution and hence infer cosmological parameters. The theory and its application was first formulated by [7]. One basic requirement of the underlying theory is to give up the geometrically thin lens approximation and look at the 3D distribution of matter which is then projected onto 2D sky. The theory of weak lensing and its applications has been reviewed many times, for a thorough review see: [].

Lets first try to model the statistical properties of the convergence field $\kappa(\theta)$. In cosmological context, the convergence field can be expressed as the weighted projection of the mass distribution integrated along the line of sight,

$$\kappa(\theta) = \int_0^{\chi_H} g(\chi)\delta(\chi\theta, \chi)\mathrm{d}\chi \tag{1.53}$$

where, δ is the 3D relative density contrast as defined in the previous section. χ_H is the comoving distance to the horizon and $g(\chi)$ is the lensing weight. Under the assumption that largest scale structure in δ are much smaller than the effective range $\Delta\chi$ of the projection (also known as the Limber's approximation), one can write the lensing weights as,

$$g(\chi) = \frac{3H_0^2\Omega_m}{2c^2}\frac{\chi}{a(\chi)\bar{n}}\int_{\chi}^{\chi_(H)} n(\chi')\frac{(\chi'-\chi)}{\chi'}\mathrm{d}\chi' \tag{1.54}$$

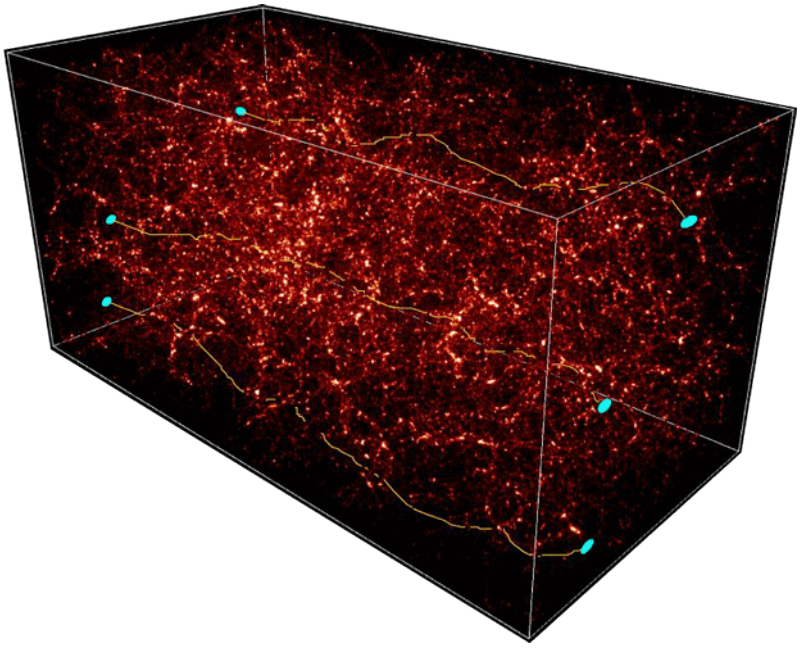

Figure 1.1.

where, $n(\chi)$ gives the distribution of sources as a function of comoving distance or redshift and $\bar{n}$ is the number of sources per unit area. These quantities depends on the experiments like Euclid [5], LSST [42] etc..

As mentioned earlier too, the density field δ is assumed to be a random field and only its statistical properties can be modelled, not the individual realisations. In the previous chapter, we modelled the second order statistics of this random field as the matter power spectrum, it is interesting to model similar quantity in lensing, the power spectrum of the convergence field and relate it to the matter power spectrum.

$$P_\kappa(\ell) \equiv \langle|\tilde{\kappa}|^2\rangle = \int_0^{\chi_H} \frac{g(\chi)^2}{\chi^2} P\left(k = \frac{\ell}{\chi}, \chi\right) d\chi \tag{1.55}$$

where, $\ell = 180/\theta$, is the multipole and give the angle in the sky.

If the power spectrum of the convergence field $P_\kappa(\ell)$ is observable, it can be used to constrain the 3D matter power spectrum and hence the cosmological parameters. Further, the same quantity can also be calculated in different redshift bins (instead of just one) and auto and cross spectra can be obtained. This process is known as lensing *tomography* and it has extra constraining power over cosmological parameters [27, 78].

Now the last problem is to relate the $P_\kappa(\ell)$ to something that is observable, possibly the cosmic shear. This is rather direct, as in the complex plane Fourier transform of the comic shear and that of the convergence can be related with a phase,

$$\tilde{\gamma(\ell)} = \exp(2i\beta)\tilde{\kappa(\ell)} \tag{1.56}$$

and therefore we have,

$$\langle|\tilde{\gamma}(\ell)|^2\rangle = \langle|\tilde{\kappa}(\ell)|^2\rangle = P_\kappa(\ell) \tag{1.57}$$

i.e., the power spectrum of the cosmic shear is the same as the cosmic shear of the convergence field. Therefore, cosmic shear can be measured in a wide survey and the power spectrum is calculated which can be used in order to put constraints on cosmology.

1.5. Motivation

The theory of structure formation is partially understood – at large scales or during early dark ages when the perturbations are tiny or only linear order is important. However, for small scale clustering processes, when the higher order in perturbations become important, the analytic solutions are not possible, and we rely on simulations and approximations. Also, as baryons play a vital role at small scale clustering and galaxy formation, it is also important to quantify these effects and understand various baryonic processes. The main motivation of this work is to model the distribution of matter in high dense regions like lensing clusters and in the large scale structures of the Universe. This information is useful in two ways:

First, the accurate modelling of the matter distribution in individual clusters give information about the properties of dark-matter. On large scales, dark-matter is known to be collision-less and non-interacting except for its gravitational effects, but it is important to quantify and test the hypothesis in high density regions, like at the centre of clusters. A small but finite cross section of dark-matter particles can be well tested in high density regions, this may have stronger implications for our understanding of the properties of dark-matter and the Universe at large. Also, as the baryonic processes and theory of galaxy formation is poorly understood, the central regions of the galaxy clusters can be used as laboratories to study these processes. It can be done only if the matter distribution in the individual systems is well constrained. Further, following the hierarchical structure formation, many small halos interacts and merge to form large collapsed virialised structures; the study of the distribution of matter in lensing clusters at high redshift give information about the merging stage of the cluster.

Modelling of the distribution of matter in the large scale structures of the Universe is very important in order to model the cosmological observables. Statistically, the modelling of the matter power spectrum or 2PCF is important if one wants to do cosmology because the matter power spectrum underlies many cosmological observables like Baryon Acoustic Oscillations (BAO), weak lensing, galaxy clustering, redshift space distortions etc.. Also, the modelling of the covariance matrix of the matter power spectrum is vital in order to do correct likelihood analysis and testing cosmological models. Finally baryonic physics also changes the power spectrum at small scales, if neglected, it will add biases in the cosmological parameters and mislead the interpretations.

1.5.1. Challenges

There are certain challenges in order to accurately model the mass distribution in clusters and in the Universe.

- Distribution of matter in clusters: Mass is not an observable, what we observe is the light in different frequency bands. We can derive redshift, velocity dispersions etc. from this. To relate these observables to mass distributions, we have approximations and various models which are often rich in systematics, i.e., the incomplete understanding and biases. Gravitational lensing is far the most unbiased technique in order to trace matter in lenses. However, even in lensing the inversion of the lens is a highly degenerate process. Often and very much in practice, people assume that mass follows light, which is a good approximation but may not be true everywhere. So, this is essential that the mass reconstruction techniques are independent of such assumptions in order to learn about intrinsic properties of the dark-matter and baryonic process in the high dense regions of the lensing clusters.

- Modelling matter power spectrum and its covariance matrix: In order to model the matter power spectrum at non-linear scales, simulations are by far the best solution as all other analytic approaches are difficult. But this is limited by the volume and resolution of the simulation. Also a good simulation can be very expensive computationally. So, it is possible to simulate a big volume at very good resolution for a cosmological realization, but in order to carry out a likelihood analysis on some cosmological data, simulations are very expensive and often impossible to do. So, we rely on semi-analytic models, which can model the matter power spectrum with some function of cosmology and are more accurate than the perturbation theories etc. at non-linear scales. Similarly, in order to get good covariance matrix of the matter power spectrum from simulation, we need to simulate a volume of the order 1000 $(h^{-1}\mathrm{Gpc})^3$ cube, which again is very expensive specially when small scale modelling is necessary.

- Modelling baryonic physics in two-point functions: Finally baryonic effects are important at small scales statistically. If these effects are not present in the model of the power spectrum, it will bias the whole exercise and the recovered cosmology might be very precise but not accurate.

The main goal of this work is to target these challenges. I performed and completed a number of projects in order to achieve these goals, As a result, I successfully completed five scientific papers that are attached with this manuscript.

Bibliography

[1] Agertz O., Teyssier R., Moore B., 2011, MNRAS, 410, 1391

[2] Ahn C. P. et al., 2012, ApJS, 203, 21

[3] Alam S. et al., 2015, ArXiv e-prints

[4] Albrecht A. et al., 2006, ArXiv Astrophysics e-prints

[5] Amendola L. et al., 2013, Living Reviews in Relativity, 16, 6

[6] Bassett B. A., Tsujikawa S., Wands D., 2006, Reviews of Modern Physics, 78, 537

[7] Blandford R. D., Saust A. B., Brainerd T. G., Villumsen J. V., 1991, MNRAS, 251, 600

[8] Blot L., Corasaniti P. S., Alimi J.-M., Reverdy V., Rasera Y., 2015, MNRAS, 446, 1756

[9] Branch D., Miller D. L., 1993, ApJ, 405, L5

[10] Brandenberger R. H., 2001, ArXiv High Energy Physics - Phenomenology e-prints

[11] Cen R., Miralda-Escudé J., Ostriker J. P., Rauch M., 1994, ApJ, 437, L9

[12] Charbonneau P., 1995, ApJS, 101, 309

[13] Chevallier M., Polarski D., 2001, International Journal of Modern Physics D, 10, 213

[14] Cooray A., Sheth R., 2002, Phys. Rep., 372, 1

[15] Croft R. A. C., Weinberg D. H., Katz N., Hernquist L., 1998, ApJ, 495, 44

[16] Davis T. M., James J. B., Schmidt B. P., Kim A. G., 2007, in American Institute of Physics Conference Series, Vol. 924, The Multicolored Landscape of Compact Objects and Their Explosive Origins, di Salvo T., Israel G. L., Piersant L., Burderi L., Matt G., Tornambe A., Menna M. T., eds., pp. 330–335

[17] Durrer R., 1994, Fund. Cosmic Phys., 15, 209

[18] Fields B., Sarkar S., 2006, ArXiv Astrophysics e-prints

[19] Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, ApJ, 473, 576

[20] Governato F. et al., 2004, ApJ, 607, 688

[21] Guedes J., Callegari S., Madau P., Mayer L., 2011, ApJ, 742, 76

[22] Hamilton A. J. S., Rimes C. D., Scoccimarro R., 2006, MNRAS, 371, 1188

[23] Heitmann K., Higdon D., White M., Habib S., Williams B. J., Lawrence E., Wagner C., 2009, ApJ, 705, 156

[24] Heitmann K., Lawrence E., Kwan J., Habib S., Higdon D., 2014, ApJ, 780, 111

[25] Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, ApJ, 715, 104

[26] Hewitt J. N., Turner E. L., Schneider D. P., Burke B. F., Langston G. I., 1988, Nature, 333, 537

[27] Hu W., 1999, ApJ, 522, L21

[28] Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

[29] Kaiser N., Squires G., 1993, ApJ, 404, 441

[30] Kiessling A., Taylor A. N., Heavens A. F., 2011, MNRAS, 416, 1045

[31] Kneib J.-P., Bonnet H., Golse G., Sand D., Jullo E., Marshall P., 2011, LENSTOOL: A Gravitational Lensing Software for Modeling Mass Distribution of Galaxies and Clusters (strong and weak regime). Astrophysics Source Code Library

[32] Komatsu E., 2006, Progress of Theoretical Physics Supplement, 163, 185

[33] Kroupa P., 2012, PASA, 29, 395

[34] Lawrence E., Heitmann K., White M., Higdon D., Wagner C., Habib S., Williams B., 2010, ApJ, 713, 1322

[35] Lee J., Pen U.-L., 2008, ApJ, 686, L1

[36] Li Y., Hu W., Takada M., 2014, Phys. Rev. D, 90, 103530

[37] Liesenborgs J., De Rijcke S., Dejonghe H., 2006, MNRAS, 367, 1209

[38] Liesenborgs J., De Rijcke S., Dejonghe H., Bekaert P., 2007, MNRAS, 380, 1729

[39] Liesenborgs J., De Rijcke S., Dejonghe H., Bekaert P., 2009, MNRAS, 397, 341

[40] Linder E. V., 2003, Physical Review Letters, 90, 091301

[41] Lotz J. et al., 2014, in American Astronomical Society Meeting Abstracts, Vol. 223, American Astronomical Society Meeting Abstracts 223, p. 254.01

[42] LSST Science Collaboration et al., 2009, ArXiv e-prints

[43] Ma C., Fry J. N., 2000, ApJ, 543, 503

[44] Marinacci F., Pakmor R., Springel V., 2014, MNRAS, 437, 1750

[45] McClelland J., Silk J., 1977, ApJ, 217, 331

[46] McQuinn M., Oh S. P., Faucher-Giguère C.-A., 2011, ApJ, 743, 82

[47] Meiksin A., White M., 1999, MNRAS, 308, 1179

[48] Meylan G., Jetzer P., North P., Schneider P., Kochanek C. S., Wambsganss J., eds., 2006, Gravitational Lensing: Strong, Weak and Micro

[49] Mohammed I., Liesenborgs J., Saha P., Williams L. L. R., 2014, MNRAS, 439, 2651

[50] Mohammed I., Martizzi D., Teyssier R., Amara A., 2014, ArXiv e-prints

[51] Mohammed I., Saha P., Liesenborgs J., 2015, PASJ, 67, 21

[52] Mohammed I., Seljak U., 2014, MNRAS, 445, 3382

[53] Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

[54] Oppenheimer B. D., Davé R., Katz N., Kollmeier J. A., Weinberg D. H., 2012, MNRAS, 420, 829

[55] Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144

[56] Perlmutter S. et al., 1999, ApJ, 517, 565

[57] Petters A. O., Levine H., Wambsganss J., 2001, Singularity theory and gravitational lensing

[58] Planck Collaboration et al., 2015, ArXiv e-prints

[59] Planck Collaboration et al., 2011, A&A, 536, A19

[60] Planck Collaboration et al., 2015, ArXiv e-prints

[61] Pontzen A. et al., 2008, MNRAS, 390, 1349

[62] Rahmati A., Pawlik A. H., Rai M., Schaye J., 2013, MNRAS, 430, 2427

[63] Refregier A., Amara A., Kitching T. D., Rassat A., Scaramella R., Weller J., Euclid Imaging Consortium f. t., 2010, ArXiv e-prints

[64] Refsdal S., 2004, in Thinking, Observing and Mining the Universe, Miele G., Longo G., eds., pp. 231–235

[65] Saha P., Williams L. L. R., 2011, PixeLens: A Portable Modeler of Lensed Quasars. Astrophysics Source Code Library

[66] Schaye J. et al., 2015, MNRAS, 446, 521

[67] Schmidt B. P. et al., 1998, ApJ, 507, 46

[68] Schneider A. et al., 2015, ArXiv e-prints

[69] Schneider P., Ehlers J., Falco E. E., 1992, Gravitational Lenses

[70] Schrodinger E., 1939, Physica, 6, 899

[71] Seljak U., 2000, MNRAS, 318, 203

[72] Semboloni E., van Waerbeke L., Heymans C., Hamana T., Colombi S., White M., Mellier Y., 2007, MNRAS, 375, L6

[73] Skillman S. W., Warren M. S., Turk M. J., Wechsler R. H., Holz D. E., Sutter P. M., 2014, ArXiv e-prints

[74] Smith R. E., et al., 2003, Mon. Not. Roy. Astron. Soc., 341, 1311

[75] Smoot G. F., 1999, in American Institute of Physics Conference Series, Vol. 476, 3K cosmology, Maiani L., Melchiorri F., Vittorio N., eds., pp. 1–10

[76] Stinson G. S. et al., 2012, MNRAS, 425, 1270

[77] Takada M., Hu W., 2013, Phys. Rev. D, 87, 123504

[78] Takada M., Jain B., 2004, MNRAS, 348, 897

[79] Takada M., Jain B., 2009, MNRAS, 395, 2065

[80] Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012

[81] Takahashi R. et al., 2009, ApJ, 700, 479

[82] Taylor A., Joachimi B., Kitching T., 2013, MNRAS, 432, 1928

[83] Theuns T., Leonard A., Efstathiou G., 1998, MNRAS, 297, L49

[84] Tsujikawa S., 2003, ArXiv High Energy Physics - Phenomenology e-prints

[85] Tyson J. A., Wenk R. A., Valdes F., 1990, ApJ, 349, L1

[86] Tytler D., O'Meara J. M., Suzuki N., Lubin D., 2000, Physica Scripta Volume T, 85, 12

[87] van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649

[88] van Daalen M. P., Schaye J., McCarthy I. G., Booth C. M., Dalla Vecchia C., 2014, MNRAS, 440, 2997

[89] Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

[90] Walsh D., Carswell R. F., Weymann R. J., 1979, Nature, 279, 381

[91] Weymann R. J. et al., 1980, Nature, 285, 641

[92] White M., Hu W., 2000, ApJ, 537, 1

2

Paper1: Massgalaxy offsets in lensing clusters

In the concordance model of cosmology, cold dark matter (CDM) provides the major budget for the matter content in the Universe, nearly 80%. It interacts only gravitationally and provides a large potential well that attracts low mass halos to form high mass halos through a series of mergers. Baryons follow these potential wells and cool down to form stars and galaxies. Therefore, it is safe to assume that light-follows-mass in these giant halos. However, there could be a number of astrophysical and/or cosmological scenarios where light-follows-mass does not hold. For example, if the dark matter interacts with baryons or with itself. This kind of non-standard properties of the dark-matter can be tested in dense regions like the centre of the galaxy clusters. Strong gravitational lensing is an ideal tool to obtain this information as it is sensitive to the total mass of the lens, and does not differentiate between dark-matter or luminous galaxies.

In this paper, we studied the mass-galaxy offsets in three lensing clusters of galaxies: Abell 3827, Abell 2218 and Abell 1689. These three clusters are very different in their morphology, total mass, redshift, and lensing data. We used GRALE, a strong gravitational lens inversion library, to model the mass maps of these clusters given the position of the multiple images of the background sources and their redshift. No information from the lensing clusters was used, except for their redshift. The mass models are completely form free. We also provide the uncertainty maps that show high-confidence in the region where lensing images were present.

Because of the free-form and high certainties of the mass maps, it is possible to compare the distribution of matter in these clusters with luminous galaxies completely on the basis of statistical uncertainties. In Abell 3827 and Abell 2218 we found

small offsets between the local mass peaks and the position of a nearby galaxy. Particularly, in Abell 3827 this offset is nearly 6 kpc, and is statistically significant. We also discussed the possible origin of these offsets, which can be either astrophysical like dynamical friction or the non-standard properties of the dark-matter component like self-interactions. The offset in Abell 3827 is further studied by Massey et al. 2015 with new HST data where the offset was found to be robust. With a simplified model, we argued that to explain this offset, the cross-section of dark-matter particles σ/m must be of the order 10^{-4}.

In Abell 1689, no significant offsets were found. However, we found a line of sight sub-structure behind the cluster at redshift ~ 3.

Role: Under the supervision of Dr. Prasenjit Saha, I started this project by modifying some parts of GRALE library for an optimised fitness function, and resolution scheme using two simulated cluster lenses: (i) a simple circular lens, and (ii) a lens with similar morphology as that of Abell 3827. I proposed the idea of reconstructing three clusters, two of which are known for mass-galaxy offsets in previous studies, and a third one without it. The motivation was to set examples of clusters having large offsets, small offsets, and no offset, which we found in Abell 3827, Abell 2218 and Abell 1689 respectively. I accumulated the lensing data for the three clusters, and reconstructed their mass maps using GRALE. In Abell 3827, I quantified the offsets statistically by locating the centre of mass in a circle around the galaxy closest to the local sub-peak. I also proposed the idea to find a line of sight substructure, if any, by using sources in bins of redshift.

This paper has been published in *Monthly Notices of the Royal Astronomical Society* (MNRAS).
Arxiv: `http://arxiv.org/abs/1402.4217`

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY **MNRAS Advance Access published February 20, 2014**

MNRAS (2014) doi:10.1093/mnras/stu124

Mass–galaxy offsets in Abell 3827, 2218 and 1689: intrinsic properties or line-of-sight substructures?

Irshad Mohammed,[1]★ Jori Liesenborgs,[2] Prasenjit Saha[1] and Liliya L. R. Williams[3]

[1]*Institute for Theoretical Physics, University of Zürich, CH-8057 Zürich, Switzerland*
[2]*Expertisecentrum voor Digitale Media, Universiteit Hasselt, Wetenschapspark 2, B-3590, Diepenbeek, Belgium*
[3]*School of Physics & Astronomy, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA*

Accepted 2014 January 15. Received 2014 January 8; in original form 2013 July 31

ABSTRACT

We have made mass maps of three strong-lensing clusters, Abell 3827, Abell 2218 and Abell 1689, in order to test for mass–light offsets. The technique used is GRALE, which enables lens reconstruction with minimal assumptions, and specifically with no information about the cluster light being given. In the first two of these clusters, we find local mass peaks in the central regions that are displaced from the nearby galaxies by a few to several kpc. These offsets *could* be due to line-of-sight structure unrelated to the clusters, but that is very unlikely, given the typical levels of chance line-of-sight coincidences in Λ cold dark matter simulations – for Abell 3827 and Abell 2218 the offsets appear to be intrinsic. In the case of Abell 1689, we see no significant offsets in the central region, but we do detect a possible line-of-sight structure: it appears only when sources at $z \gtrsim 3$ are used for reconstructing the mass. We discuss possible origins of the mass–galaxy offsets in Abell 3827 and Abell 2218: these include pure gravitational effects like dynamical friction, but also non-standard mechanisms like self-interacting dark matter.

Key words: gravitational lensing: strong – galaxies: clusters: individual: Abell 1689 – galaxies: clusters: individual: Abell 2218 – galaxies: clusters: individual: Abell 3827 – galaxies: haloes.

1 INTRODUCTION

Our current understanding of the Universe and its dynamics indicate that its major components are dark: cold dark matter (CDM) and the so-called dark-energy. Unlike baryons, dark matter interacts only gravitationally and provides the deep potential wells which are followed by the baryons. The baryons form clumps at these potential wells and cool down to form stars. The standard ΛCDM model explains a range of observed processes pretty well, from the angular power spectrum of the cosmic microwave background (Planck Collaboration et al. 2013) to the baryonic acoustic oscillations (Sánchez et al. 2013) in the large-scale structure and the number counts of clusters. However, the intrinsic properties and behaviour of dark matter and dark energy remain an open problem in cosmology.

In the picture of hierarchical structure formation in ΛCDM model, galaxy clusters are the most recently formed structures that are gravitationally bound. They are cosmic laboratories to test the laws of gravity, structure formations and the interaction of different species of particles. A galaxy cluster contains lots of galaxies – tens to thousands, hot intracluster plasma visible in X-rays, a variety of relativistic particles and finally dark matter which dominates its mass budget. Measuring the mass of the galaxy cluster is an essential aspect of using the cluster to study many other things. There are several physical processes that enable one to measure the mass: the kinematics of cluster galaxies (Saro et al. 2013), the hydrodynamics of hot gas emitting X-rays (Vikhlinin et al. 2009), and gravitational lensing. Lensing is particular interesting, because it relies only on gravity and does not itself require any luminous objects in the cluster being studied. One of the questions that lensing can address is how well the luminous matter traces the distribution of total mass. Deviations, or lack thereof, from the mass-follows-light hypothesis will provide important information about the physical processes going in within clusters. The first lensing-based detection of deviations from mass-follows-light goes back to the late 1990s (Abdelsalam, Saha & Williams 1998) but the observation that generated a wide interest in these deviations was that of the Bullet Cluster (Clowe et al. 2006), which showed unambiguously that dark matter is quite collisionless compared to the gas phase baryonic matter (Randall et al. 2008). While the properties of dark matter are probably not the only reason for deviations from mass-follows-light in galaxy clusters, dark matter self-interaction cross-section and how to optimally extract it from observations is an exciting avenue of research (Harvey et al. 2013a,b).

This work uses strong gravitational lensing to look for deviations from mass-follows-light, i.e. it explores the correspondence on the sky between the dark-matter peaks with the galaxies in the

★ E-mail: irshad@physik.uzh.ch

central parts of three galaxy clusters, Abell 3827, 2218 and 1689. These clusters are very different from each other in morphology and redshift. As we discuss in Section 4, some deviations we find may be due to the non-standard properties of dark matter, but others could be the result of superimposed substructure, or hydrodynamics within the cluster.

We use GRALE (Liesenborgs, De Rijcke & Dejonghe 2006; Liesenborgs et al. 2007), a strong-gravitational lensing tool to reconstruct the mass map of the clusters. There is no overall parametric form for the mass distribution, but rather an adaptive grid. Other than the redshift, no information about the cluster is required as input, not even its location or morphology. This makes GRALE well suited to reconstruction of mass maps before comparison with light.

2 THE LENS-RECONSTRUCTION TECHNIQUE

GRALE has been applied to other strong-lensing clusters (Liesenborgs et al. 2008, 2009) and compared with other techniques (Zitrin et al. 2010, 2011), so here we just give a general description and then some tests.

2.1 GRALE

The data given to GRALE consist of the identified multiple-image systems and their redshifts, along with possible regions where additional images are guessed to be likely. No information about the light from the lens is given. The mass maps in GRALE are free-form, being made up of a superposition of many components. In this work, each component is taken as a Plummer lens, that is, the usual Plummer sphere

$$\rho = \frac{3M}{4\pi}\frac{a^2}{(r^2+a^2)^{5/2}} \tag{1}$$

projected to two dimensions. Other choices of lens component, such as square tiles, are also possible.

Any mass distribution in GRALE is assigned a fitness with respect to the given data. The fitness has two components, as follows.

(i) For a given mass map, the input images are ray-traced back to the source, using the lens equation. The more nearly these back-projected images coincide for any multiple-image system, the fitter the mass map. If the fitness measure were simply the source-plane distance between the back-projected images, that would favour extreme magnification (tiny sources); accordingly, the fitness measure is scaled to the source size.

(ii) There could be further places in the image plane that, when ray-traced back to the source, coincide with the sources corresponding to the observed images. These correspond to extra images, and would be favoured by the above fitness measure. There may indeed be undiscovered extra images in certain regions, but in most of the image plane, extra images can be ruled out with high confidence. The area of no images present is referred to in GRALE as the null space. For each image system, the user specifies a null space, which is simply the image plane with the images themselves cut out, and (optionally) further cut-outs where incipient images could potentially be present. Images in the null space lead to a fitness penalty for the mass map.

It is possible to have other components to the fitness, such as time delays for quasar source (Liesenborgs et al. 2009), but this work uses these two. The null space, item (ii) above, is a unique aspect of GRALE. There are other techniques that allow the mass distribution to be very general in form, as with GRALE, but they make additional assumptions in order to suppress extra images, such as constraining local density gradients (Saha, Read & Williams 2006) or applying smooth interpolation schemes Coe et al. (2008). Only GRALE incorporates the absence of images as useful data.

The computational part of GRALE is optimizing the fitness function for the given data, using a genetic algorithm. The basic idea, inspired by Darwinian evolution, is to generate a population of trial solutions. A fitness measure is assigned to each trial solution and then these solutions are combined, cloned and mutated to get the next generation of populations supported by a better fitness function. Genetic algorithms have long been used in astrophysics for hard optimization problems (for a somewhat old but readable review, see Charbonneau 1995). They tend to be computationally expensive, but are often effective on otherwise intractable problems. GRALE uses a multi-objective genetic algorithm, meaning that the different components of the fitness function are compared individually, not just combined into a single function. Only the fitness ranking matters in genetic algorithms, not the actual values of the fitness. In terms of likelihoods and posterior probabilities, models with better fitness are considered more probable, that is, the fitness components are monotonic in the posterior probability, but there is no known or assumed functional relation between likelihood and fitness.

The locations and masses of the Plummer components are chosen by the genetic algorithm. The algorithm also adapts the number of Plummers, but an allowed range is specified by the user. That is, the user specifies the level of substructure. For the GRALE fitness measure, lower is better, and it decreases as we increase the resolution of the map. This is quite intuitive as more Plummer spheres naturally result in a better fit. So the overall criterion should be somehow a function of the GRALE fitness measure and the number of Plummers. We are not aware of any theoretical argument that yields the appropriate criterion, but after some experimentation we found one that works reasonably well in test cases. This is an 'unfitness' or

$$\text{badness} = \ln\left(\text{GRALE fitness} \times \sqrt{\text{number of components}}\right). \tag{2}$$

If we think of the GRALE fitness measure as a mismatch distance, and the number of Plummers as the inverse resolution length, the badness criterion appears natural.

To choose the number of Plummer components, we adopted the following procedure. First, we have GRALE reconstruct the lens with a comparatively low number of Plummers. Then, we let GRALE improve the fit with progressively more Plummers, allowing more substructure to be introduced. After that, we let GRALE continue to adapt the fit with progressively fewer Plummers. The mass distribution with the minimum badness (equation 2) is taken as the result.

We now report on two simulated lenses, which we generated and then reconstructed with GRALE, in order to check the pipeline and calibrate the error estimates.

2.2 A simple lens

A Plummer lens of mass $10^{14}\,M_{\odot}$ was generated at redshift 0.1. Six sources were put at different redshifts (one at 0.15, two at 0.2, two at 0.4 and one at 1.0). The mass profile and image plane are shown in Fig. 1. The images and source redshifts were given to

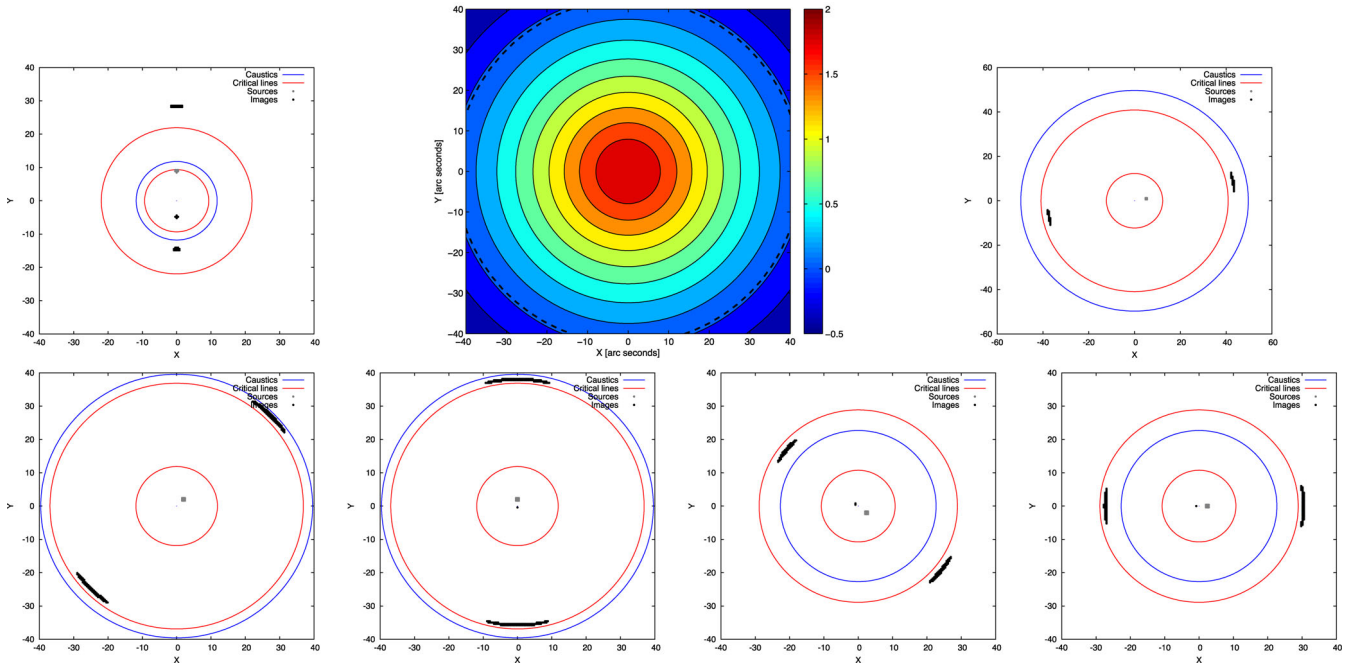

Figure 1. A circularly symmetric synthetic lens (centre top panel) and six image systems from sources at different redshifts. Sources are in grey, caustics are in blue, critical curves are in red. The contour lines in the synthetic lens are those of constant surface mass density; the colour scale is in units of log (kg m^{-2}). The same scale is used in all figures in this paper. For reference, $\Sigma_{\rm crit}$ for $z_{\rm l} = 0.1$ and $z_{\rm s} = 0.2$ in a standard ΛCDM cosmology is 18.7 kg m^{-2}.

the inversion module of GRALE. Fig. 2 shows the reconstructed masses at different resolutions and the badness values.

When reconstructing the lens, GRALE did not have the information that in fact it had a simple parametric form, without substructures. The reconstructions do have some substructure, as well as small offsets from the centre. Such spurious features increase with resolution. The least-badness criterion, however, favours a model with relatively little substructure.

2.3 A more complex lens

We now increase the complexity, both of the input lens and of the reconstruction procedure. For each data set, from now on we will present a mean map Σ and a fraction rms-deviation map $\delta\Sigma/\Sigma$, obtained as follows. From the images, we first let GRALE construct a sequence of maps at nine different resolutions (as with the simple lens), and then select the one at minimum badness. This whole procedure is repeated 10 times, to obtain an ensemble of reconstructions. The mean and rms deviation refer to such an ensemble, as

$$\delta\Sigma = (\langle\Sigma^2\rangle - \langle\Sigma\rangle^2)^{1/2}\,. \tag{3}$$

Each map of Σ and $\delta\Sigma/\Sigma$ comes out of 90 separate reconstructions at different resolutions. The typical computational requirement is 50 h $\times$ 16 cores.

A simulated lens at redshift 0.1 was next created with five Plummers positioned such that the configuration resembles the inner region of Abell 3827. Sources were put at different redshifts, as follows.

(i) Three-source case: three sources at $z = 0.2$ were given as input.
(ii) Four-source case: a fourth source at $z = 0.4$ was added.
(iii) Five-source cases: a fifth source at $z = 1.0$ was added.

The resulting images, along with caustics and critical curves, are shown in Fig. 3). Results from these are shown in Fig. 4. The top row of the figure shows the mass maps Σ. The second row shows $\delta\Sigma/\Sigma$, or the fractional rms deviation. The third row shows $\Delta\Sigma/\delta\Sigma$ where $\Delta\Sigma$ is the (absolute) actual deviation of the reconstructed mass map from the real mass map. If $\delta\Sigma$ were close to $\Delta\Sigma$, we could simply take the rms deviation as the uncertainty. In fact the rms deviation underestimates the true error by about a factor of 2. That can be read off the bottom row of Fig. 4, which plots the cumulative distribution of $\Delta\Sigma/\delta\Sigma$.

The main result from this test is that the rms deviation times two is a reasonable approximation of the errors. In addition, we can also read off some qualitative features from Fig. 4. First, the spur or handle-like feature to the lower right is recovered in the lens reconstruction in all cases, even if not perfectly reproduced. Secondly, the maps get more accurate as more sources, especially at different redshifts, are introduced.

We conclude that GRALE is able to find offsets as well as extended structures (if any) in lenses.

3 RECONSTRUCTION OF THREE REAL CLUSTERS

In this section, we do mass reconstructions of three galaxy-clusters, and present these with their accompanying mass error maps. The two sets of maps for each cluster allow us to judge whether light-follows-mass (LFM) is a good assumption. We defer the discussion of the implications of the deviations from LFM to Section 4.

3.1 Abell 3827

Abell 3827 is a lensing cluster at redshift 0.099. Three multiply lensed image systems have been identified (Carrasco et al. 2010) belonging to three sources at redshift 0.204, most probably different

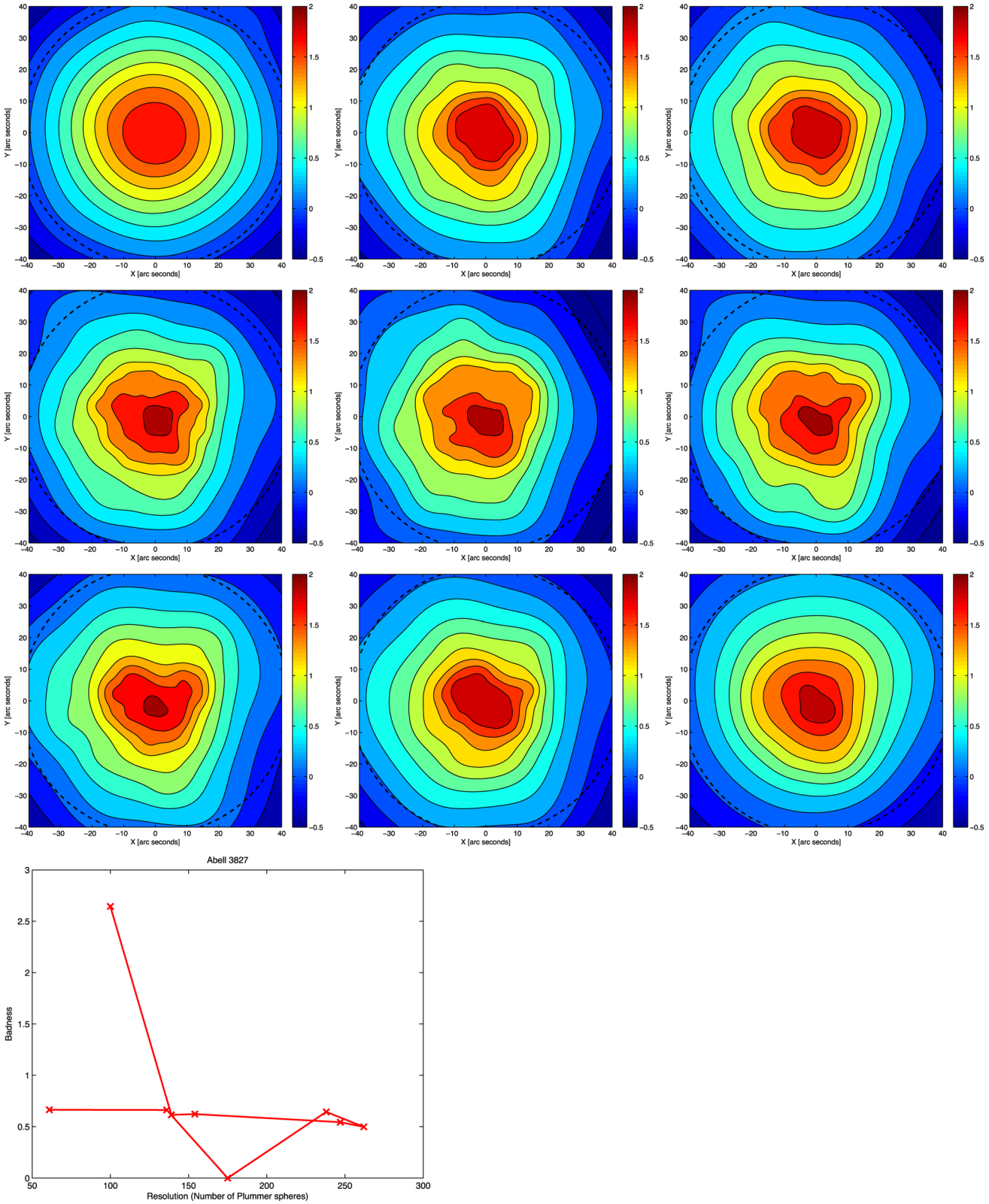

Figure 2. Reconstruction of the lens in Fig. 1 from the data in that figure. The badness curve (bottom panel) shows that the best model is the third one (top-right map in the grid of nine.) The dashed circle in each map delineates the modelled region . The sequence of mass maps is in reading order (from top left to bottom right).

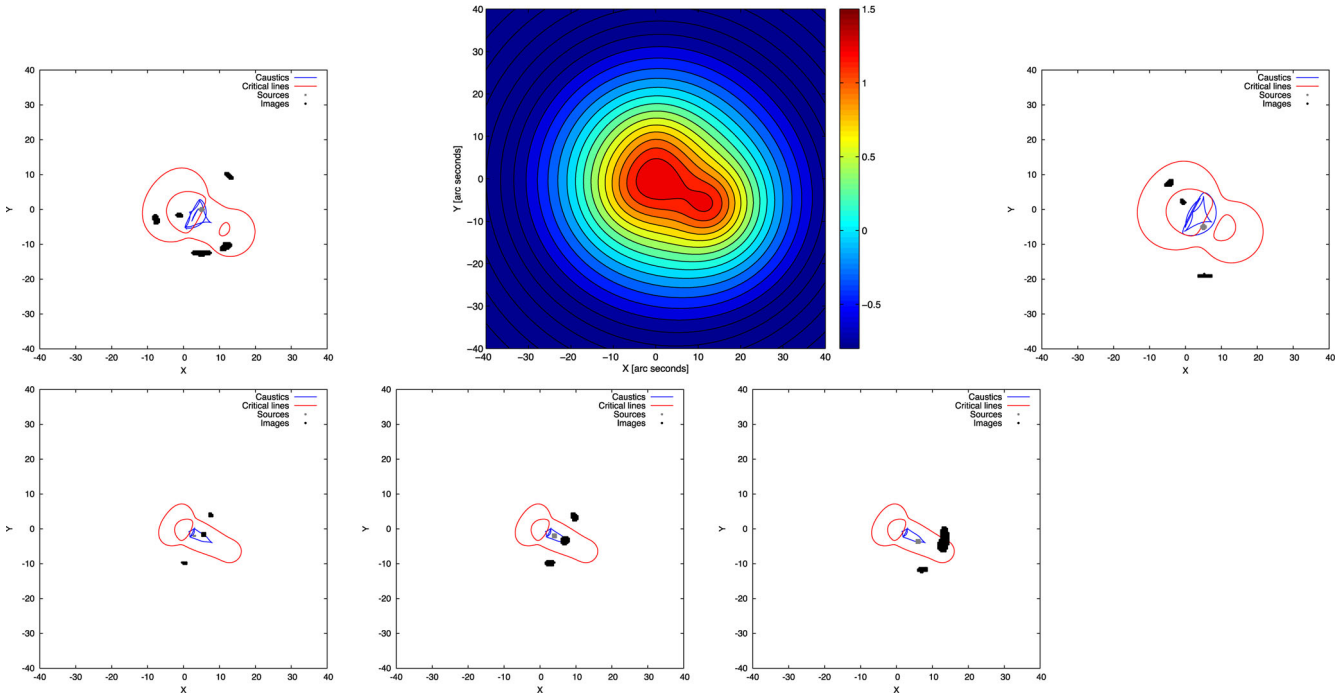

Figure 3. A synthetic lens with a main mass concentration and a nearby secondary mass peak. Five projected Plummer spheres are used to construct this lens. Image systems from five sources at different redshifts are shown in separate panels.

parts of the same source. Another big arc is identified belonging to a source at redshift 0.408, but its multiply imaged counterpart has not yet been identified. A mass map based on these images (Williams & Saha 2011) indicates a dark extended clump, offset by ∼6 kpc from the brightest of the four or five ellipticals in the cluster core. This offset, if confirmed, would afford us a unique opportunity to examine and understand the dynamics in dense regions of clusters. One of the primary goals of this paper is to assess the reality of this offset and estimate its statistical significance. GRALE is a very different lens mass reconstruction method from the one used in Williams & Saha (2011), so detecting the offset with GRALE will lend credence to its reality.

Using the identified images we reconstructed the mass distribution in two ways, and then combined the results. These are displayed in the three rows of Fig. 5.

First, we used the three image systems belonging to the sources at redshift 0.2. The first panel of the top row of Fig. 5 shows a spur in the mass map, which is offset from the nearby elliptical galaxy (the rightmost of the five grey dots). The spur's location is similar to the location of the local overdensity reported in Williams & Saha (2011), so the offset is similar in both reconstructions. From the map of fractional rms deviation $\delta\Sigma/\Sigma$ (right-hand panel of the first row) the spur appears to be significant; the rms deviation in that region is about 0.1 kg m^{-2}, and so the fractional error is about 10 per cent. Since the structure appears to be extended and not a single clump, it is not obvious how to quantify it. We can nonetheless test its significance. We chose a circle of radius 5 arcsec (green circle) around the nearby elliptical. (The choice of size is somewhat arbitrary; other choices would also serve our purpose.) We then calculate the centre of mass within this circle, for each mass map within the ensemble, and mark them with green '+' signs in the middle panel of top row, which is a zoom on to the relevant region. All 10 centroids are consistently displaced from the nearby galaxy (grey circle), by about 1.2 arcsec. The average of the 10 centroids is marked with a blue star symbol. We may interpret these results as a hypothesis test. The null hypothesis is that the cluster has no mass/galaxy offset, and the mass is centred on the galaxy light. A mass reconstruction could nonetheless put the aperture centroid displaced from the galaxy, simply from the stochastic element in the genetic algorithm – note that the mass reconstructions are not given any information about the cluster galaxies. If there is no mass offset, the model offsets would be random, and the change of all 10 mass reconstructions having an offset in the same direction would be only 10 per cent. But the aperture centroids are consistently offset in the same region. Hence, there does appear to be an offset, significant at 90 per cent confidence, between the mass spur and the galaxy.

Secondly, we used all four image systems: three belonging to the sources at redshift 0.2 and one with source redshift 0.4. As mentioned before, no image counterpart of the latter has been identified, but there is a possibility of such a counter-image near the centre of the cluster. Accordingly, we allowed GRALE to produce extra images in that region. The corresponding mass maps are shown in the second row of Fig. 5. This time the extent of the image region is larger, and the fraction rms between reconstructions (right-hand panel) is smaller in the general region of the image at $z_l = 0.4$. A clear mass subpeak is seen near the elliptical, offset from it by ∼4 arcsec or ∼7 kpc. To be consistent with the previous case, we again calculate the centre of mass, or centroid, in a circular region of radius 5 arcsec. Individual centroids are marked with green '×' signs, and their average is the blue star. Again the offset is detected at a significance similar to the one above.

Finally, we then combined the two sets of ensembles described above, for a total of 20 individual maps. The bottom row of Fig. 5 shows the average mass map, and the map of $\delta\Sigma/\Sigma$ for the combined ensemble. The conclusion remains unchanged.

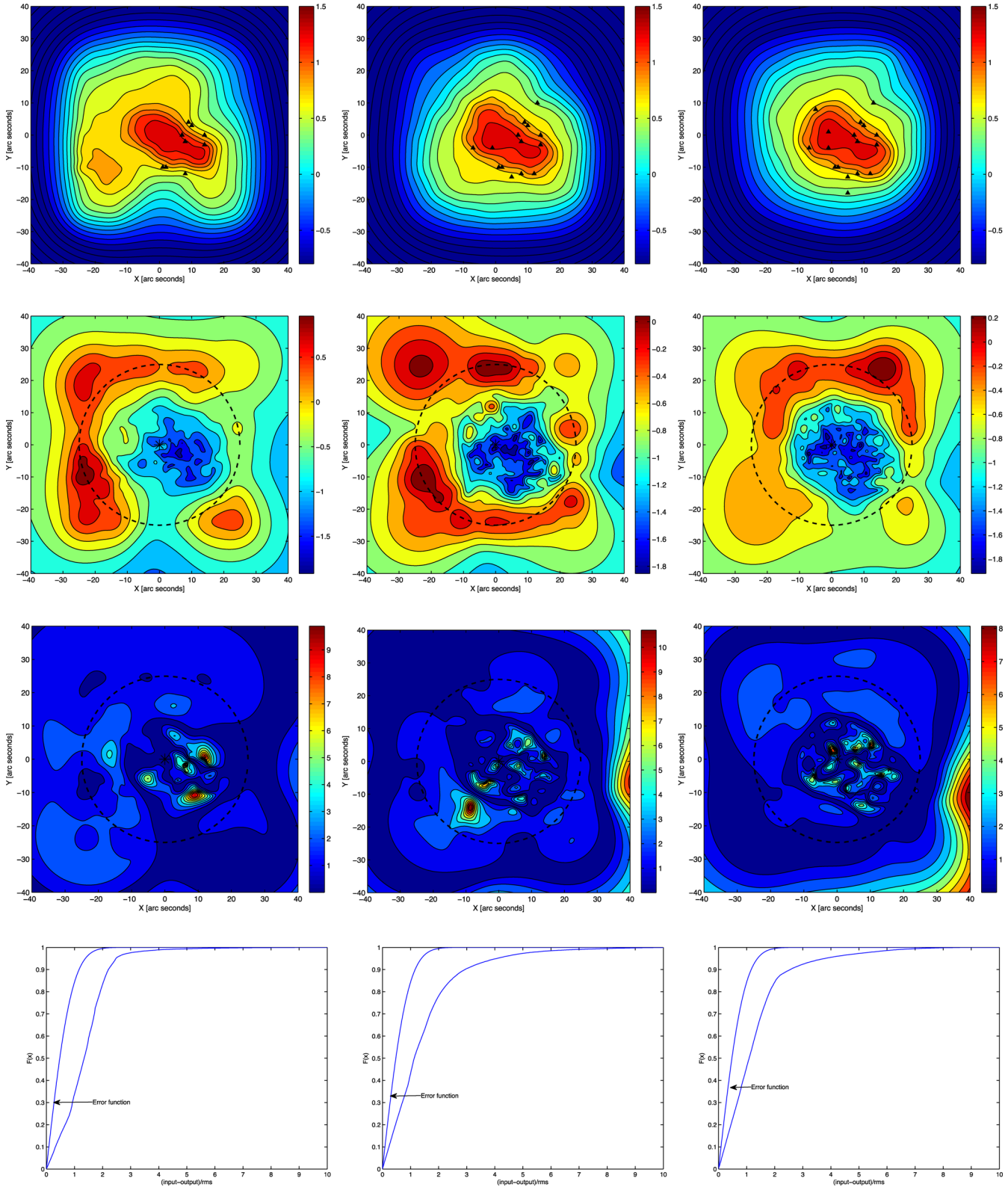

Figure 4. Reconstruction of the lens in Fig. 3. Column 1: using three sources only, with the corresponding images shown as black triangles; column 2: using four sources; column 3: using all five sources. The top row shows average surface mass density Σ; units are same as in Fig. 1. The second row shows the fractional rms deviation of 10 reconstructions, $\delta\Sigma/\Sigma$. The third row contains $\Delta\Sigma/\delta\Sigma$, where $\Delta\Sigma$ is the pixel-wise difference between the true map and the average reconstructed map. The bottom row shows the cumulative $\Delta\Sigma/\delta\Sigma$, along with the corresponding curve (marked 'error function') for Gaussian errors with dispersion $\delta\Sigma$. We conclude that the error estimate $\delta\Sigma$ needs to be multiplied by $\sim$2 (or increased by 0.30 on a $\log_{10}$ scale). The worst cases are some very small regions (red in the lower panels) where $\log_{10}\Delta\Sigma$ should be increased by $\sim+1$.

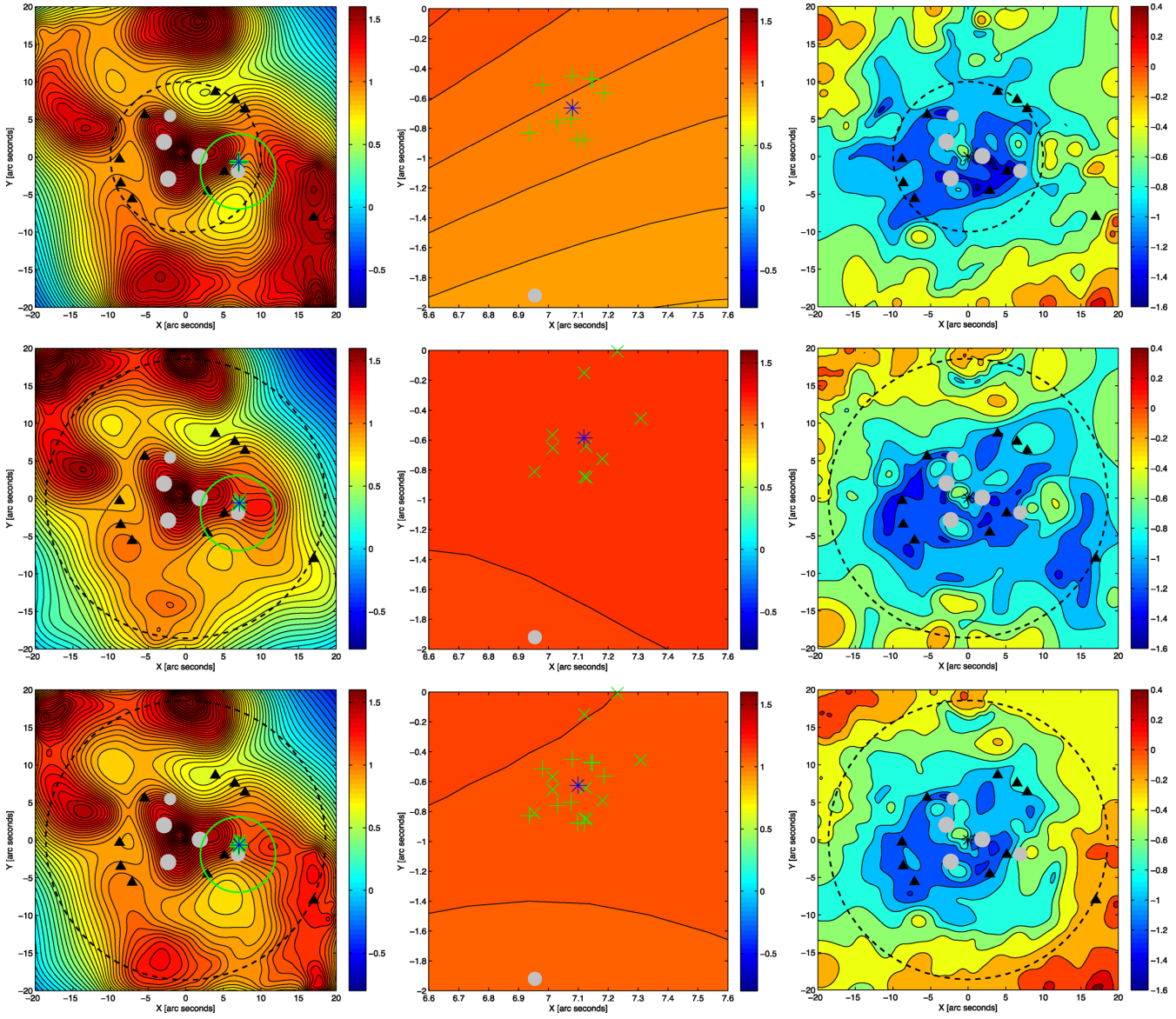

Figure 5. Mass reconstructions of A3827. North is up and east to the right. The scale is 1.82 kpc arcsec^{-1}. The upper row maps are an ensemble of 10 maps, each obtained using only the 9 images of the source at $z_s = 0.2$. The middle row shows an ensemble of 10 maps, using 9 images of the $z_s = 0.2$ source and the single image at $z_s = 0.4$. The bottom row combines both ensembles. The left-hand column presents the average of the 10 mass maps. The middle column is a zoom centred on the most luminous elliptical N1. The 10 green '+' signs (top row) and '×' signs (middle row) represent centroids from 10 individual maps of the mass within the green circle shown in the left-hand column. The grey dot towards the bottom of the plots (in the middle column) is N1. The blue asterisk is the centroid of the average of the 10 realizations. The right-hand column shows the fractional rms deviation between the 10 maps, $\delta\Sigma/\Sigma$.

3.2 Abell 2218

Abell 2218 is a well known and much studied lensing cluster (e.g. Abdelsalam et al. 1998). Like other rich clusters, it has been used in the recent years as a cosmic telescope (Altieri et al. 2010; Hopwood et al. 2010; Knudsen et al. 2010) to get a better view of distant or faint galaxies. The strong-lensing region is somewhat larger on the sky than in Abell 3827, and the greater redshift, $z_l = 0.175$, implies a larger physical scale, 3 kpc arcsec^{-1}.

We reconstructed the cluster using the four most secure strong-lensing systems. Fig. 6 shows the mass map (left-hand panel) and fraction rms dispersion between the 10 individual maps of the ensemble (right-hand panel). While apparent offsets are visible between galaxies (grey dots) and mass in the central region of the cluster, these are not significant, because rms in that region is comparable to the typical value of the surface mass density. Significant offsets are seen around the lower-right mass clump, where the rms dispersion between mass maps is low. In the central panel, we show a zoom of that region, similar to that in the middle panel of Fig. 5. The green '+' signs represent the local mass peaks (not centroids as in the case of A3827) of individual reconstructions, which are displaced from the nearest cluster galaxies, represented by grey dots in the upper right of that panel.

3.3 Abell 1689

Abell 1689, at redshift 0.183, is perhaps the best-known lensing cluster, containing over a hundred lensed images from at least 30 background sources extending to high redshifts (Broadhurst et al.

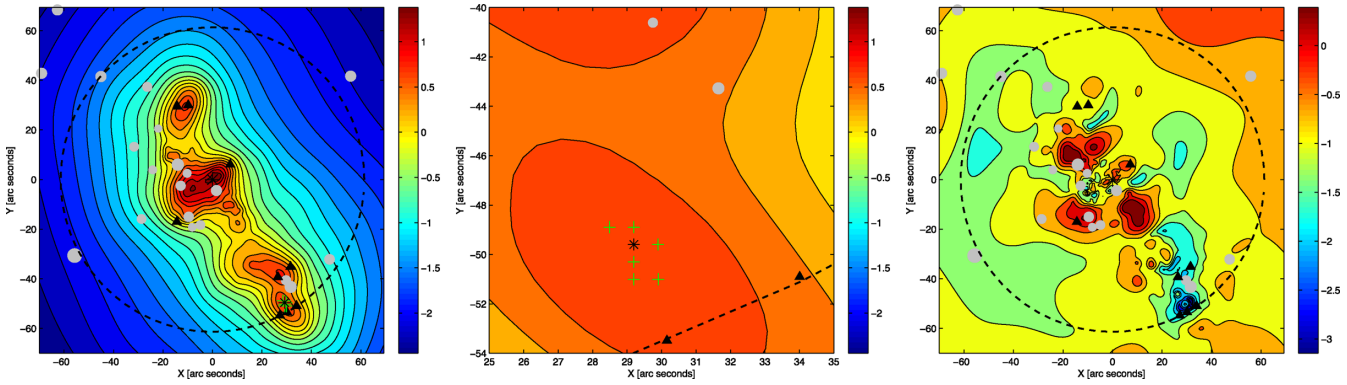

Figure 6. Mass map of A2218. North is up and east to the right. The average mass map (left-hand column) and fractional rms (right-hand column) are based on 10 realizations. The central column shows the zoom of the region with mass–light offsets, and the green '+' signs are the local mass peaks from individual reconstructions. The scale is 3 kpc arcsec^{-1}. Galaxies with $R < 20$ (Pello et al. 1992) are marked with grey dots.

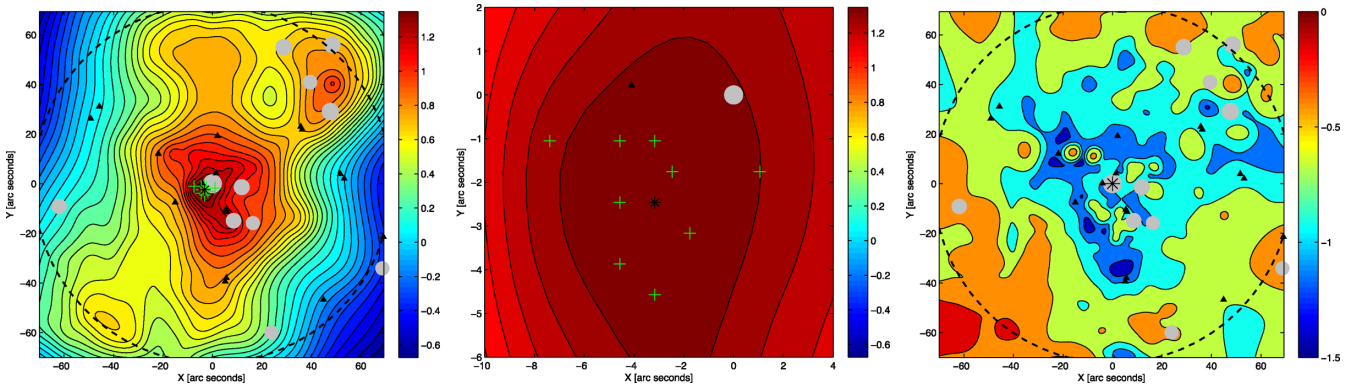

Figure 7. Mass maps of A1689. North is up and east to the right. The columns are similar to those in Fig. 6. Galaxy positions (Duc et al. 2002) also marked.

2005). Our reconstruction of its mass is shown in Fig. 7. As with Abell 2218, the mass map and the rms maps are in the left- and right-hand panels. There are no significant mass/light offsets in this cluster. To illustrate that, in the central panel we show a zoom into the central region, where the mass peaks of the 10 individual maps are shown as green '+' symbols. Their distribution with respect to the central cluster galaxy (grey dot) is consistent with the two being coincident.

Because the cluster has many multiply imaged systems spanning a wide range of redshifts, it is possible to test if there are line of sight (los) structures that have affected the positions of images. We divided the multiply lensed sources into two groups, the low-redshift system (LRS) and high-redshift system (HRS). LRS consists of a total of three multiply imaged systems with five, three and three (total of eleven) images at redshifts 2.54, 1.99 and 1.98, respectively. HRS consists of a total of two multiply imaged systems with two and five (total of seven) images at redshifts 4.53 and 2.99, respectively. We then carried out mass reconstruction for A1689 using LRS and HRS separately. The two mass maps are shown in Fig. 8, in the upper-left and upper-middle panels, respectively. The corresponding fraction rms distributions are shown below each map. The upper-right panel is the difference between HRS and LRS maps divided by the rms of the LRS maps ($\Delta\Sigma/\delta\Sigma$). Most of this map is consistent with a uniform surface mass density of low amplitude, about a factor of 10 below the critical surface mass density. This could be due to steepness, or mass sheet degeneracy which affected one map more than the other. The only prominent feature is a mass excess in the HRS map, compared to the LRS map, centred at around (-20 arcsec, 35 arcsec). The $\delta\Sigma$ maps for both HRS and LRS are both low in that region, suggesting that the structure is real. We interpret this feature as a los structure, probably in the redshift range 2–3. Another test of the structure's significance is shown in the lower right, which contains a histogram of the upper-right plot $\Delta\Sigma/\delta\Sigma$ (pixel wise). The putative los structure contributes to the tail extending beyond the right edge of the distribution. The corresponding lensing mass would be $\sim 10^{13}\,\mathrm{M}_{\odot}$ if the structure were at the same redshift at A1689, but since the structure can only be at $z > 2.5$, the critical density and hence the lensing mass are much lower – a few times $10^{12}\,\mathrm{M}_{\odot}$ – amounting to a modest galaxy group. There is another feature at (-50 arcsec, -60 arcsec), but it is outside the image circle, and the $\delta\Sigma$ in that region says that it is not significant.

4 DISCUSSION

Gravitational lensing offers a unique opportunity to study the distribution of matter in clusters of galaxies. Free-form reconstruction methods take full advantage of this. Our synthetic tests show that GRALE recovers the mass distribution well, and the concomitant errors provide a reliable guide to assessing the significance of various mass features. The test case in Figs 2 and 4 shows no spurious offsets in the mass maps.

Reconstructions of the three real lensing clusters indicate some curious features. In two clusters we see offsets between the optical light and the nearest mass concentrations. The form of the offsets is not resolved: they could be distinct peaks in the projected mass

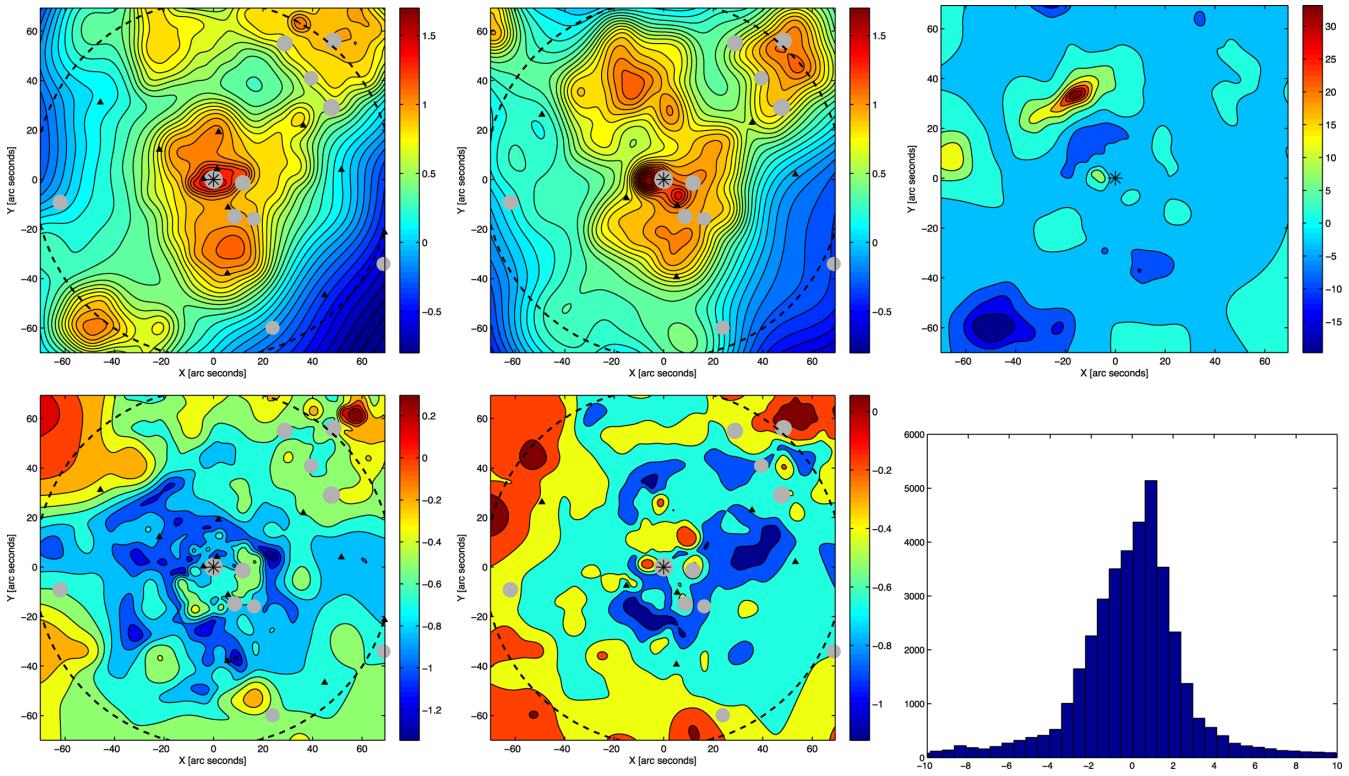

Figure 8. Test for the los structure in A1689. Upper-left and upper-middle panels are the mass maps obtained using two separate sets of sources: at low and high redshifts, respectively. Lower-left and lower-middle panels are the corresponding fractional rms maps. Upper right is the difference between the high-z (HRS) and the low-z (LRS) maps divided by the rms of the low-z maps (i.e. $\Delta\Sigma/\delta\Sigma$, which is dimensionless); the scale is linear. Note the apparent structure at higher z, near (-20 arcsec, 35 arcsec). Lower right is the histogram of the map above it (pixel wise) $\Delta\Sigma/\delta\Sigma$.

distribution; or they could be spurs that extend from a peak that itself coincides with the galaxy light; or the offsets could very lopsided dark haloes around galaxies. (We emphasize that not all offsets seen in the reconstructed mass maps are significant, but only those that pass the statistical significance tests.) A caveat to bear in mind is the assumption that the observed image positions are accurate. Because lensed images are often faint, have low surface brightness and are superimposed on brighter cluster galaxies, image identification is not always straightforward. It is thus conceivable that some images have been misidentified. But assuming that the image identifications are all valid, confirmation by independent techniques is desirable. Lens reconstruction methods not assuming light traces mass in some way include Lensview (Wayth & Webster 2006), LensPerfect (Coe et al. 2008) and PBL (Deb, Goldberg & Ramdass 2008) and any of these would be suitable. If the mass/galaxy offsets are confirmed, they would lead to interesting conclusions about the nature of clusters and dark matter.

In general, several reasons for offsets are possible. Superimposed, but dynamically unrelated los structures could contribute lensing mass, with no apparent associated light, especially if the structures are considerably further away from us than the main lensing cluster. However, we argue that the offset in A3827, is not due to the los structure because of the very low redshifts of the sources. In A2218, los structures are also unlikely to be the cause because only a very concentrated and massive los structure can contribute significantly in the vicinity of a massive clump within a cluster. Such chance superposition are expected to be rare.

Line-of-sight structures are more likely to make a contribution away from mass concentrations within the cluster, where cluster projected densities are lower. This can be illustrated with dark-matter N-body simulations. The blue lines in Fig. 9 are the isodensity contours of the total projected mass in a cylinder centred on a halo whose virial radius is the radius of the window, while the red lines are the contours of the projected mass inside the virial sphere of the cluster. We caution that these plots were made with a limited los depth of about comoving 90 Mpc (Simulations courtesy Jürg Diemand; Diemand, Moore & Stadel 2004). The black contours mark regions where the fractional mass excess due to the los structures (and not the mass within the virial sphere) amount to 25 per cent of total. The top two panels show examples where the contribution from the los material is typical, while the bottom two panels present two cases with the most contribution (out of a total of 100 lines of sight). Even though the length of the cylinder is not large, the plots show that los structures cannot make a significant contribution where the cluster density is high. However, such structures can make a significant contribution at some distance away from the cluster centre.

In A1689, we might be seeing such an los structure. After subtracting the mass reconstruction based on HRS from that based on LRS we see a mass concentration about 30 arcsec, or 100 kpc from cluster centre. It is statistically significant (it contributes to the tail of the distribution shown in Fig. 8 which extend beyond the right edge of the plot) but is not associated with bright cluster galaxies. We interpret it as arising from a structures between the $z \approx 2$ and 3.

If not los structure, what else can be responsible for mass–light offsets seen in A3827 and A2218? Offsets could be intrinsic to the cluster, and be due to manifestations of known physics, like gravity and hydrodynamics of the gas, or new physics, such as self-scattering of dark matter. Offsets in merging clusters have been observed, but mostly between the dark matter and the X-ray

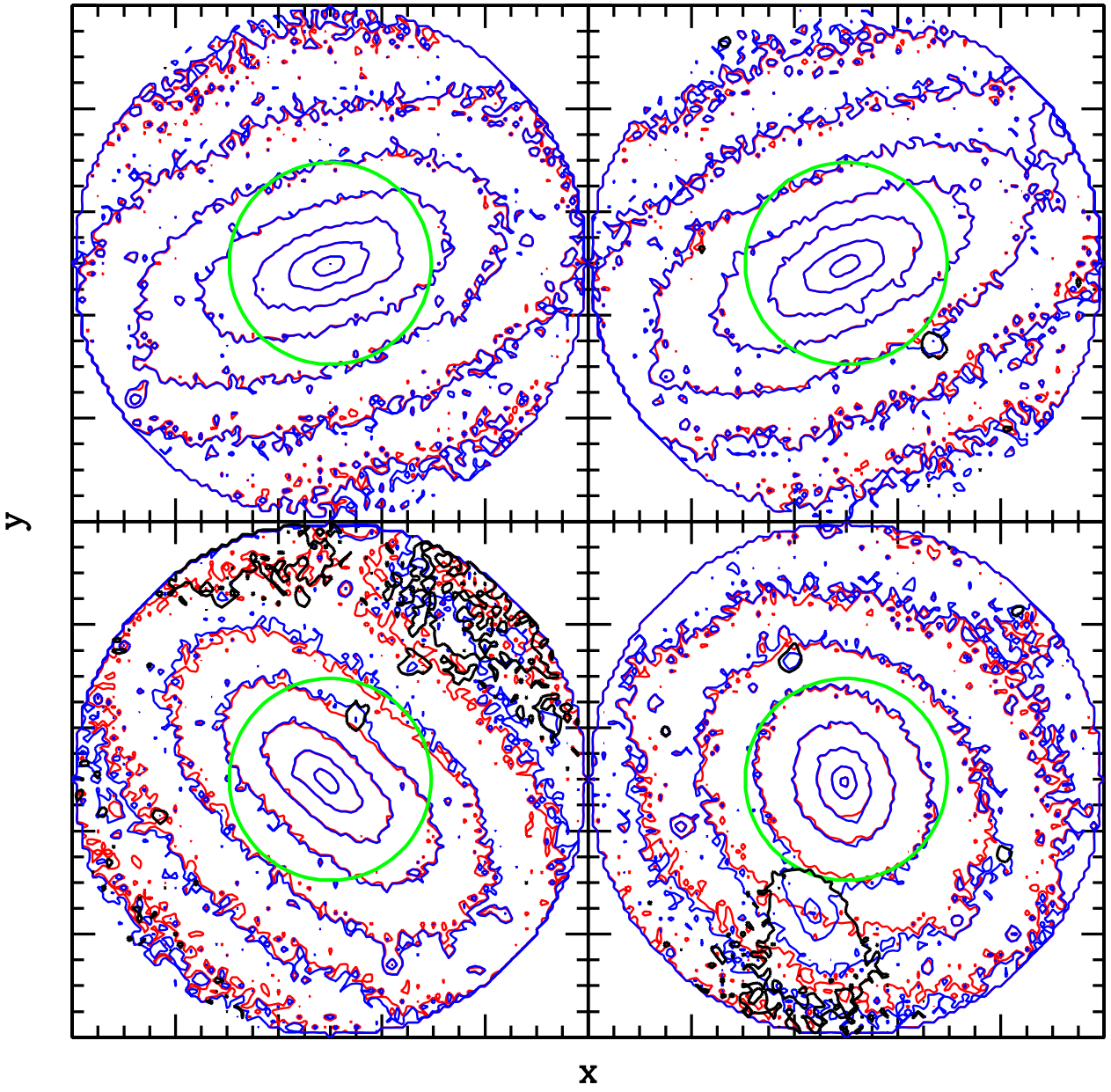

Figure 9. Density contours of projected mass centred on haloes taken from dark-matter-only simulations (Diemand et al. 2004). The radius of the window is the virial radius, and the green circle marks the typical radius where lensed images will be formed. The red density contours are due to the halo mass interior to the virial sphere, while the blue contours are due to all projected mass within a cylinder of roughly 90 Mpc. The black contours mark regions where the fractional mass excess due to the los structures (and not the mass within the virial sphere) amount to 25 per cent of total. The top two panels show average lines of sight, while the bottom panels the two (out of 100) where los material makes the most contribution.

emitting gas components (Clowe et al. 2006, 2012). In the outskirts of Abell 2744, a separation between dark matter and galaxy components is also seen (Merten et al. 2011), and in the merging cluster CL0152-1357 an offset between Sunyaev–Zel'dovich effect and X-ray peaks has been detected (Molnar, Hearn & Stadel 2012). Most of these offsets are on larger scales then what we detect in this work. For smaller scale offsets, early, stage mergers are probably not the cause, and different set of causes has to be considered.

One of the possibly relevant gravitational effects is the oscillation or wobbling of a galaxy, such as a BCG around the bottom of the gravitational potential. This has been observed in a sample of galaxy clusters as a displacement of the BCG from the lensing centroid (Zitrin et al. 2012). The distribution is displacements is wide, and peaks at roughly 10 kpc. Whether this is a likely explanation for the offsets in A3827 and A2218 is yet to be determined – the observed offsets are not for central cluster galaxies.

It is less likely, but still possible that the offsets are a consequence of tidal effects. These would strip the material from the galaxy symmetrically in the leading and trailing directions. Since the offsets in A3827 and A2218 do not show such symmetry, tidal effects are probably not the main cause.

Dynamical friction would create an asymmetric structure and would preferentially distort the distribution of dark matter and not stars if the former has a more extended distribution. A numerical simulation would be required to test this possibility.

The formation of a galaxy cluster is a complex process involving hydrodynamics of gas. It is possible that star formation induced by galaxy mergers within clusters would result in stars and dark-matter haloes offsets.

Finally, if dark matter has non-negligible self-interaction cross-section, dark-matter particles of the galaxy halo would experience a drag force as the galaxy moves within the halo of the cluster. The nature of the resulting dark matter features induced by these interactions may be consistent with those observed in A3827 and A2218, but detailed simulations are required (Kahlhoefer et al. 2014).

ACKNOWLEDGEMENTS

LLRW would like to acknowledge the hospitality of ITP, Zurich.

REFERENCES

Abdelsalam H. M., Saha P., Williams L. L. R., 1998, AJ, 116, 1541
Ade P. A. R. et al. (Planck Collaboration), 2013, preprint (arXiv:1303.5075)
Altieri B. et al., 2010, A&A, 518, L17
Broadhurst T. et al., 2005, ApJ, 621, 53
Carrasco E. R. et al., 2010, ApJ, 715, L160
Charbonneau P., 1995, ApJS, 101, 309
Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109
Clowe D., Markevitch M., Bradač M., Gonzalez A. H., Chung S. M., Massey R., Zaritsky D., 2012, ApJ, 758, 128
Coe D., Fuselier E., Benítez N., Broadhurst T., Frye B., Ford H., 2008, ApJ, 681, 814
Deb S., Goldberg D. M., Ramdass V. J., 2008, ApJ, 687, 39
Diemand J., Moore B., Stadel J., 2004, MNRAS, 352, 535
Duc P.-A., Poggianti B. M., Fadda D., Elbaz D., Cesarsky C., 2002, A&A, 382, 60
Harvey D., Massey R., Kitching T., Taylor A., Jullo E., Kneib J.-P., Tittley E., Marshall P. J., 2013a, MNRAS, 433, 1517
Harvey D. et al., 2013b, preprint (arXiv:1310.1731)
Hopwood R. et al., 2010, ApJ, 716, L45
Hsu L.-Y., Ebeling H., Richard J., 2013, MNRAS, 429, 833
Kahlhoefer F., Schmidt-Hoberg K., Frandsen M. T., Sarkar S., 2014, MNRAS, 437, 2865
Knudsen K. K., Kneib J.-P., Richard J., Petitpas G., Egami E., 2010, ApJ, 709, 210
Liesenborgs J., De Rijcke S., Dejonghe H., 2006, MNRAS, 367, 1209
Liesenborgs J., De Rijcke S., Dejonghe H., Bekaert P., 2007, MNRAS, 380, 1729
Liesenborgs J., De Rijcke S., Dejonghe H., Bekaert P., 2008, MNRAS, 389, 415
Liesenborgs J., De Rijcke S., Dejonghe H., Bekaert P., 2009, MNRAS, 397, 341
Merten J. et al., 2011, MNRAS, 417, 333
Molnar S. M., Hearn N. C., Stadel J. G., 2012, ApJ, 748, 45
Pello R., Le Borgne J. F., Sanahuja B., Mathez G., Fort B., 1992, A&A, 266, 6
Randall S. W., Markevitch M., Clowe D., Gonzalez A. H., Bradač M., 2008, ApJ, 679, 1173
Saha P., Read J. I., Williams L. L. R., 2006, ApJ, 652, L5
Sánchez A. G. et al., 2013, MNRAS, 433, 1202
Saro A., Mohr J. J., Bazin G., Dolag K., 2013, ApJ, 772, 47
Vikhlinin A. et al., 2009, ApJ, 692, 1033
Wayth R. B., Webster R. L., 2006, MNRAS, 372, 1187
Williams L. L. R., Saha P., 2011, MNRAS, 415, 448
Zitrin A. et al., 2010, MNRAS, 408, 1916
Zitrin A., Broadhurst T., Coe D., Liesenborgs J., Benítez N., Rephaeli Y., Ford H., Umetsu K., 2011, MNRAS, 413, 1753
Zitrin A., Bartelmann M., Umetsu K., Oguri M., Broadhurst T., 2012, MNRAS, 426, 2944

This paper has been typeset from a TEX/LATEX file prepared by the author.

3

Paper2: Analytic model for the matter power spectrum

The future generation cosmological surveys, like Euclid, LSST etc., are expected to provide very large quantity of high quality data such that it will be possible to probe small scale structures like never before. Employing the cosmological observables from these experiments, like weak gravitational lensing, baryon acoustic oscillations etc., percent level constraints are expected on all cosmological parameters. With such tight constraints, either the ΛCDM model will gain more credibility in the community else it would be rejected with higher confidence; either way the efforts will be unified. In order to employ the full constraining power from these experiments, a good understanding of non-linear structure formation is needed. In this paper, we studied an important aspect of this theory – the matter power spectrum. It underlies many cosmological observables, it is important to model it accurately and precisely up to the non-linear regime.

In this paper, we provide an estimator for the matter power spectrum based on the Zeldovich approximation and the halo model. This model is calibrated on $N-$body simulations, and gives an accuracy of a few percent up to $k \sim 0.8h^{-1}$Mpc over a range of cosmological models including neutrino masses and redshifts.

We also provide an estimator for the full covariance matrix of the matter power spectrum which is very important for statistical inference from the cosmological data. In spite of the simple form of the covariance estimator, it is found to be in remarkable agreement with simulations.

We provide a description of baryonic effects on the matter power spectrum in this framework. This model can be used to project and estimate weak lensing power spectrum and utilise the future generation surveys to put strong and unbiased con-

straints on the cosmological parameters.

Role: This project was done under the supervision of Professor Uros Seljak. The starting point was analytical calculations using the halo model. The idea was to expand the 1-halo term in Taylor series, and analytically evaluate the coefficients. I accumulated 38 matter power spectra from the cosmic emulator, each for its original cosmological node. I also evaluated Zeldovich power spectra using code from Zvonimir Vlah for the same 38 cosmological models and three redshifts (0.0,0.5,1.0). I fitted the function $A_0 - A_2k^2 + A_4k^4$ to the difference between the full matter power spectra and Zeldovich approximation, and compared the fitted coefficients with those evaluated analytically from the halo model. We found big differences in A_2 and A_4 coefficients, whereas A_0 remains the same. This shows the divergence of the halo model from the true matter power spectrum in non-linear scales. I found very strong correlation between fitted coefficients and σ_8, and therefore we fitted them using a single power law for all three redshifts and computed the corresponding residuals. Further I fitted the residuals for its correlation to the effective slope $n_{\rm eff}$. We provide these fitting functions in the paper to estimate the coefficients of the 1-halo term using the cosmological models and redshifts. Given these fitting functions, we computed the matter power spectra, and compared them to the original emulator output. I found an agreement of about a percent up to $k \sim 0.7h^{-1}$Mpc. I accumulated the power spectra from van Daalen et al. 2011, for both dark-matter only as well as hydrodynamical simulations, and to their difference, I fitted the similar function. The only alteration was that this time I recovered the change in the coefficients due to the baryonic effects. We found that A_0 is indifferent to the baryonic effects, while A_2 and A_4 change significantly, implying the conservation of mass inside the halo where the profile is changing due to baryons. In the the same vein, we argued that at scales where only Zeldovich term is important, the covariance is dominated by the cosmic variance. To compute the total covariance, we can compute the variance of each of the coefficients. However, due to the baryonic effects contaminating A_2 and A_4, it is good to add the variance of A_0 only to the total covariance, and marginalise over the other two parameters. This form of the covariance matrix is found to be in remarkable agreement with the simulations.

This paper has been published in *Monthly Notices of the Royal Astronomical Society* (MNRAS).
Arxiv: `http://arxiv.org/abs/1407.0060`

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **445,** 3382–3400 (2014) doi:10.1093/mnras/stu1972

Analytic model for the matter power spectrum, its covariance matrix and baryonic effects

Irshad Mohammed[1]★ and Uroš Seljak[2]

[1]*Physik-Institut, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
[2]*Department of Physics, Department of Astronomy and Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA*

Accepted 2014 September 19. Received 2014 September 19; in original form 2014 June 30

ABSTRACT

We develop a model for the matter power spectrum as the sum of Zeldovich approximation and even powers of k, i.e. $A_0 - A_2k^2 + A_4k^4 - \ldots$, compensated at low k. With terms up to k^4, the model can predict the true power spectrum to a few per cent accuracy up to $k \sim 0.7\,h\,\mathrm{Mpc}^{-1}$, over a wide range of redshifts and models. The A_n coefficients contain information about cosmology, in particular amplitude of fluctuations. We write a simple form of the covariance matrix as a sum of Gaussian part and A_0 variance, which reproduces the simulations remarkably well. In contrast, we show that one needs an N-body simulation volume of more than 1000 $(\mathrm{Gpc}\,h^{-1})^3$ to converge to 1 per cent accuracy on covariance matrix. We investigate the supersample variance effect and show it can be modelled as an additional parameter that can be determined from the data. This allows a determination of σ_8 amplitude to about 0.2 per cent for a survey volume of $1(\mathrm{Gpc}\,h^{-1})^3$, compared to 0.4 per cent otherwise. We explore the sensitivity of these coefficients to baryonic effects using hydrodynamic simulations of van Daalen et al. We find that because of baryons redistributing matter inside haloes all the coefficients A_{2n} for $n > 0$ are strongly affected by baryonic effects, while A_0 remains almost unchanged, a consequence of halo mass conservation. Our results suggest that observations such as weak lensing power spectrum can be effectively marginalized over the baryonic effects, while still preserving the bulk of the cosmological information contained in A_0 and Zeldovich terms.

Key words: neutrinos – methods: analytical – galaxies: haloes – galaxies: statistics – cosmological parameters – large scale structures of Universe.

1 INTRODUCTION

The clustering of dark matter (DM) as a function of scale and redshift contains useful information about many cosmological parameters. For example, clustering as a function of redshift is very sensitive to the dark energy density and its equation of state. Clustering as a function of scale can reveal information about the primordial slope of the power spectrum and matter density, as well as about the presence of massive neutrinos. The best way to measure the DM clustering is via weak lensing (Bartelmann & Schneider 2001; Refregier 2003). In weak lensing light from distant galaxies, called sources, is being deflected by mass distribution along the line of sight, such as that the images are distorted. The primary distortion is shear, which changes ellipticity of the light of the source galaxy. By correlating these ellipticities between the source galaxies one can deduce the clustering strength of the matter along the line of sight. Over the past decade this recognition put weak lensing surveys at the forefront of cosmological probes, with several ground based and space based experiments proposed (Hoekstra et al. 2006; Massey et al. 2007; Fu et al. 2008; Schrabback et al. 2010). The primary statistic is the convergence power spectrum $C_l^{\kappa\kappa}$, which can be expressed as a weighted projection over the matter power spectrum $P(k)$ along the line of sight from the observer to the source. Future surveys will contain sources at many different redshifts, and by combining this information one can minimize the line-of-sight projection and measure a quantity close to the three-dimensional power spectrum, a procedure called weak lensing tomography. In this paper, we will focus on the three-dimensional power spectrum of matter $P(k)$.

★ E-mail: irshad@physik.uzh.ch

The procedure to extract information from the weak lensing measurements is in principle straightforward and, while experimentally challenging, its theoretical underpinnings have been known for a long time. What are the remaining theoretical challenges in this programme? The predictions of the dark matter only (DMO) clustering on small scales, where non-linear effects are important, was one of the uncertainties. For example, the widely used HALOFIT (Smith et al. 2003) is only accurate to 10 per cent, although the revised version (Takahashi et al. 2012) is argued to be 5 per cent accurate for $k < 1.0\,h\,\mathrm{Mpc}^{-1}$. Recent progress in *N*-body simulations suggests this problem will soon be solved. For example, the *The Coyote Universe* DMO power spectrum emulator (Heitmann et al. 2009, 2010; Lawrence et al. 2010) is accurate to nearly 1 per cent up to $k \sim 1\,h\,\mathrm{Mpc}^{-1}$ for the 38 cosmologies that have been simulated. The emulator provides an output power spectrum for any cosmological model, interpolated from the grid of 38 simulated models, with an error that can be as high as 5 per cent for some cosmological models. It seems likely that the precision will reach the required level in the near future as finer grids of simulations are developed, but it is also clear that by using better ways to interpolate between the models could improve the accuracy.

The second problem are the baryonic effects. Baryons differ from the DM in several aspects. First difference is that hot baryonic gas has pressure, which prevents clustering on small scales. These effects are particularly important inside the DM haloes, where gas temperature is high and pressure effects large. In addition, baryons cool and condense into stars, possibly bringing DM along in the process. However, baryons also form stars, which in turn lead to supernovae that can produce energy outflows. Even more dramatic effects can arise from the active galactic nuclei (AGN), which can also produce massive energy outflows. Recent studies with hydrodynamical simulations (van Daalen et al. 2011) have argued that these AGN feedback models are required to match the observations of X-ray groups and clusters, specially the temperature–luminosity relation in X-rays. The outflowing baryons can also redistribute the DM. Recent work (Semboloni et al. 2011; Semboloni, Hoekstra & Schaye 2013) shows that the baryonic correction in the matter power spectrum can be important above $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ and if one does not take account for it, it will bias the cosmological constraints such as dark energy equation of state (Semboloni et al. 2011).

Third theoretical problem that remains unsolved is the issue of reliable covariance matrix for the observed power spectrum and optimal weighting of the data. The full covariance matrix consists of two parts: Gaussian and non-Gaussian. Both scale inversely with the volume of the survey. Gaussian contribution is very large at large scales (low wavemodes k) due to sample variance, i.e. finite number of long wavelength Fourier modes sampled in a finite volume. At higher k, the sampling variance becomes small, and Fisher matrix calculations based on Gaussian variance have predicted that most of the cosmological information in weak lensing comes from small scales. However, the non-Gaussian part becomes important on smaller scales and makes these predictions unreliable. There are two essential contributions to the covariance matrix: one arises from the Poisson fluctuations in the number of haloes relative to the average, and the second arises from the fluctuations on the scale of the survey, which induce curvature type effects that couple to all modes inside the survey (Baldauf et al. 2011; Takada & Hu 2013). For weak lensing applications these contributions become significant for $\sim\ell > 500$ (Yoo & Seljak 2012). So far the predictions have relied either on the halo model (Takada & Hu 2013) or on the simulations (Sato et al. 2009, 2011; Li, Hu & Takada 2014). It has been argued that large numbers of simulations are needed to converge for a single model (Sato et al. 2011; Blot et al. 2014). Without a reliable covariance matrix one cannot optimally combine the different power spectrum estimates, nor can one reliably estimate the errors, as emphasized in recent work (Percival et al. 2014; Taylor & Joachimi 2014).

In this paper, we propose a different approach to the DM power spectrum description that addresses all of the challenges above. We propose a novel form of the halo model for the DM power spectrum (Ma & Fry 2000; Peacock & Smith 2000; Seljak 2000; Cooray & Sheth 2002), in which we split the power spectrum into the quasi-linear 2-halo term, which we take to be the Zeldovich approximation (ZA), and the 1-halo term. Rather than relying on the analytic forms for the 1-halo term as in the original halo model (Seljak 2000; Ma & Fry 2000), we simply expand it into the series of even powers of k and fit each coefficient to the simulations. By doing so we obtain an accurate description of the DM power spectrum up to $k \sim 0.7\,h\,\mathrm{Mpc}^{-1}$. We then investigate the baryonic effects on these coefficients and address the question how to marginalize against these effects. Finally, the resulting solution we propose also simplifies the question of the covariance matrix calculations.

The outline of the paper is as follows: In Section 2, we review some important theoretical background, particularly the halo model (Section 2.1) and ZA (Section 2.2). We postulate the necessary modifications in the 1-halo term in Section 2.3 and calibrate the fitting functions on simulations in Section 3 and showing the comparison with the true matter power spectrum. In Section 4, we discuss the covariance matrix and cosmological information content of our model. We are also discussing supersample variance in Section 4.3. In Section 5, we describe the same method with baryons and the limits to which one can calculate the non-linear matter power spectrum and its full covariance matrix using this methodology. Finally in Section 6, we summarize and discuss the possibility of the future work.

2 THEORETICAL MODEL FOR DM POWER SPECTRUM

2.1 The halo model

There are several approaches to account for clustering of DM and its evolution in the Universe. One of the more successful frameworks is the halo model (McClelland & Silk 1977; Ma & Fry 2000; Peacock & Smith 2000; Seljak 2000; Cooray & Sheth 2002). We will first review the halo model as implemented in previous work before presenting a new version of the halo model that is more accurate. In the halo model approach, all the matter in the Universe is assumed to be in isolated haloes with mass defined by a threshold density as

$$M_{\triangle} = \frac{4}{3}\pi R_{\triangle}^3 \triangle \bar{\rho}_{\mathrm{m}}, \tag{1}$$

where M is the mass of the halo inside the radius $R_\triangle$ and the density of the halo is $\triangle$ times $\bar{\rho}_{\rm m}$, which is the mean matter density of the universe. We use $\triangle = 200$ throughout this paper unless stated otherwise. The power spectrum can be split into two parts:

$$P(k) = P_{\rm 1h}(k) + P_{\rm 2h}(k), \tag{2}$$

where the two terms in right are the 1-halo and 2-halo term, respectively. The 2-halo term gives the correlation between different haloes, also referred as halo–halo term, whereas the 1-halo term describe the correlation between DM particles within the halo, also referred to as Poisson term, and dominates at smaller scales. These two terms are given by

$$P_{\rm 1h}(k) = \int {\rm d}\nu\, f(\nu) \frac{M}{\bar{\rho}} |u(k|M)|^2, \tag{3}$$

$$P_{\rm 2h}(k) = \left[\int {\rm d}\nu\, f(\nu) b(\nu) u(k|M)\right]^2 P_{\rm L}(k), \tag{4}$$

where $P_{\rm L}(k)$ is the linear power spectrum. Throughout this paper, we use publicly available code CAMB (Lewis, Challinor & Lasenby 2000) to compute linear matter power spectrum, unless stated otherwise. We also used publicly available code CHOMP[1] to compute some functions like the halo mass function and density profiles. The Fourier transform of the density profile of the haloes $u(k|M)$ is normalized such that $u(k = 0|M) = 1$,

$$u(k|M) = \frac{4\pi}{M} \int_0^{R_{\rm vir}} {\rm d}r\; r^2\; \rho(r|M)\; \frac{\sin(kr)}{kr}. \tag{5}$$

One can see that upon expanding $\sin(kr)/kr$ only even powers of k will be present, as further developed below. The functions $f(\nu)$ and $b(\nu)$ are the mass function and halo bias, respectively. Both variables ν and M account for the scale and related as

$$\nu(M, z) = \left(\frac{\delta_c}{\sigma(M, z)}\right)^2, \tag{6}$$

where $\delta_c \sim 1.68$,

$$\sigma^2(M, z) = \sigma^2(M) D_+^2(z), \tag{7}$$

$$\sigma^2(M) = \frac{1}{2\pi^2} \int {\rm d}k\; k^2\; P_{\rm L}(k)\; |\bar{W}(kR)|^2, \tag{8}$$

with, $D_+(z)$ as the growth factor and $\bar{W}(x)$ as the Fourier transform of the top-hat function:

$$\bar{W}(x) = 3\frac{\sin(x) - x\cos(x)}{x^3}. \tag{9}$$

2.2 The new 2-halo term: ZA

The halo model is not sufficiently accurate for the 1 per cent precision required from the future surveys. The 2-halo term needs to be modified because in the halo model it is essentially given by the linear theory, and the non-linear effects such as the smearing of baryonic acoustic oscillations (BAO) are ignored. A useful improvement is the ZA (Zel'dovich 1970). In it we assume the particles stream along the initial trajectory, without being perturbed by the non-linear effects. Even though the ZA is in a sense linear, its effects on the density extend beyond linear effects, and ZA can even lead to caustics where the density is infinite. While ZA produces too little power to be a good approximation for the fully non-linear power spectrum, it smears the BAO in the amount that matches the simulations quite well (Taylor 1993; Matsubara 2008). As such it is a useful extension of the linear power spectrum. Here, we will consider ZA approximation for large scales, coupled to the 1-halo term for the small scales.

The Zeldovich power spectrum is given by (see e.g. Schneider & Bartelmann 1995)

$$(2\pi)^3 \delta^D(k) + P(k) = \int {\rm d}^3 q\; {\rm e}^{-{\rm i}\boldsymbol{q}\cdot\boldsymbol{k}} \exp\left[-\frac{1}{2} k_i k_j \mathbf{A}_{ij}(\boldsymbol{q})\right], \tag{10}$$

where

$$A_{ij}(\boldsymbol{q}) = X(q)\delta_{ij}^K + Y(q)\hat{q}_i\hat{q}_j, \tag{11}$$

and

$$X(q) = \int_0^\infty \frac{{\rm d}k}{2\pi^2} P_{\rm L}(k) \left[\frac{2}{3} - 2\frac{j_1(kq)}{kq}\right], \tag{12}$$

[1] http://code.google.com/p/chomp/

$$Y(q) = \int_0^\infty \frac{\mathrm{d}k}{2\pi^2} P_{\rm L}(k) \left[-2 j_0(kq) + 6 \frac{j_1(kq)}{kq} \right]. \tag{13}$$

Here, $P_{\rm L}(k)$ is the linear power spectrum and j_n is the spherical Bessel function of order n.

2.3 The 1-halo term expansion

In this section, we first motivate the 1-halo term expansion into even powers of k. In the next section, we analyse their dependence on the cosmological parameters and compare against the predictions of the halo model.

We begin by writing the ansatz for the 1-halo term,

$$P_{\rm 1h}(k) = (A_0 - A_2 k^2 + A_4 k^4 - \cdots) F(k). \tag{14}$$

To motivate the ansatz and calculate the coefficients A_n, we start with the Fourier transform of the normalized density profile, assuming for now $F(k) = 1$:

$$u(k|M) = \frac{4\pi}{M} \int_0^{R_{\rm vir}} \mathrm{d}r\, r^2\, \rho(r|M)\, \frac{\sin(kr)}{kr}. \tag{15}$$

The halo profile is spherically averaged and assumed to depend only on the mass of the halo. We can model the halo density profile in the NFW form (Navarro, Frenk & White 1997)

$$\rho(r|M) = \frac{\rho_{\rm s}}{(r/R_{\rm s})(1 + r/R_{\rm s})^2}. \tag{16}$$

This model assumes that the profile shape is universal in units of scale radius $R_{\rm s}$, while its characteristic density $\rho_{\rm s}$ at $R_{\rm s}$ or concentration $c = R_{\rm vir}/R_{\rm s}$ may depend on the halo mass M.

The function $\sin(kr)/kr$ can be expand as Taylor series with even powers of kr as

$$u(k|M) = \frac{4\pi}{M} \int_0^{R_{\rm vir}} \mathrm{d}r\, r^2\, \rho(r|M) \left[1 - \frac{k^2 r^2}{3!} + \frac{k^4 r^4}{5!} - \cdots \right]. \tag{17}$$

We can simplify this equation using function $\Im_n$ as

$$u(k|M) = \Im_0 k^0 - \Im_1 k^2 + \Im_2 k^4 - \cdots \equiv (-1)^n \sum_{n=0}^{\infty} \Im_n k^{2n}, \tag{18}$$

and

$$|u(k|M)|^2 = (-1)^{m+n} \sum_{(m,n)} \Im_n k^{2n} \Im_m k^{2m} = (-1)^{m+n} \sum_{(m,n)} \Im_m \Im_n k^{2(m+n)}, \tag{19}$$

where

$$\Im_n = \frac{4\pi}{(2n+1)! M} \int_0^{R_{\rm vir}} \mathrm{d}r\, r^{2(1+n)}\, \rho(r|M). \tag{20}$$

Note that the functions $\Im_n$ are the integrals over the density profiles and some power of r from 0 to $R_{\rm vir}$ and that $\Im_0 = 1$. However, there is nothing obviously special about truncating the integral there, and it can be changed to truncate the density profile at a different $R_{\rm max}$ than $R_{\rm vir}$, for example $2R_{\rm vir}$. This suggests that the halo model has some flexibility in its implementation and is not fully predictive. For this reason, we will just use it as a motivation and will not be doing the actual integrals over the halo profiles.

Next, we insert equation (19) into 1-halo term expression of equation (3) and group the terms in even powers of k,

$$P_{\rm 1h}(k) = \int \mathrm{d}\nu\, f(\nu) \frac{M}{\bar{\rho}} \sum_{(m,n)} \Im_m \Im_n k^{2(m+n)}, \tag{21}$$

$$P_{\rm 1h}(k) = \int \mathrm{d}\nu\, f(\nu) \frac{M}{\bar{\rho}} \left[\Im_0 \Im_0 k^0 - 2\Im_0 \Im_1 k^2 + (\Im_1 \Im_1 + 2\Im_0 \Im_2) k^4 - \cdots \right]. \tag{22}$$

Comparing equations (14) and (22), we obtain the coefficients and their variances as

$$\begin{aligned} A_0 &= \int \mathrm{d}\nu\, f(\nu) \frac{M}{\bar{\rho}}\, \Im_0 \Im_0 \\ A_2 &= \int \mathrm{d}\nu\, f(\nu) \frac{M}{\bar{\rho}}\, 2\Im_0 \Im_1 \\ A_4 &= \int \mathrm{d}\nu\, f(\nu) \frac{M}{\bar{\rho}}\, (\Im_1 \Im_1 + 2\Im_0 \Im_2) \end{aligned} \tag{23}$$

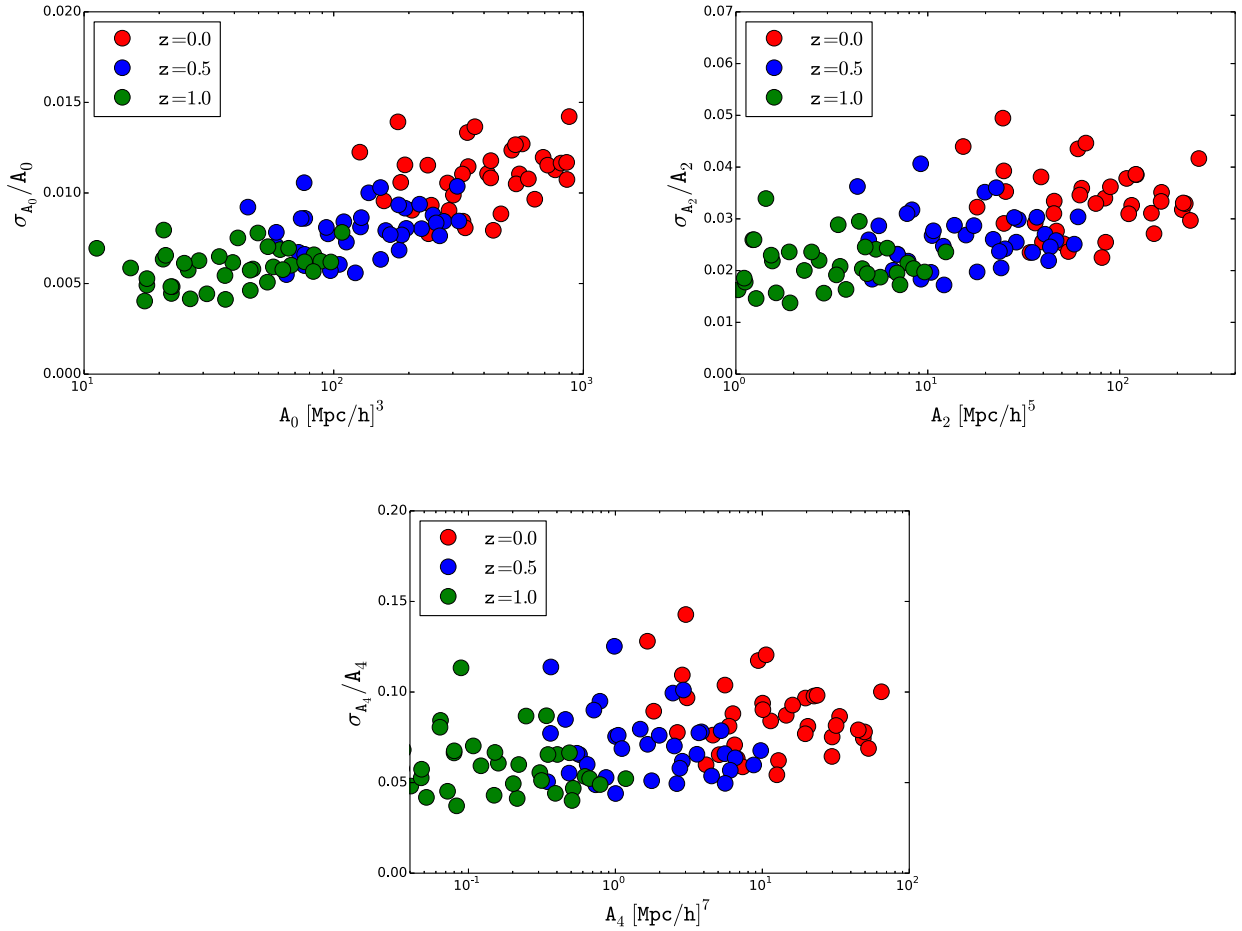

Figure 1. Relative variance $\Delta A_{2n}/A_{2n}$ versus A_n based on our model for A_0, A_2 and A_4 for three different redshift: 0.0 (red), 0.5 (blue) and 1.0 (green). Each circle bullet is one cosmological realization of the 38 cosmic emulator nodes.

with covariance,

$$\mathbf{Cov}(A_i A_j) = \begin{pmatrix} \int d\nu\, g(\nu)\,(\mathfrak{I}_0\mathfrak{I}_0)^2 & \int d\nu\, g(\nu)\,(\mathfrak{I}_0\mathfrak{I}_0)(2\mathfrak{I}_0\mathfrak{I}_1) & \int d\nu\, g(\nu)\,(\mathfrak{I}_0\mathfrak{I}_0)(\mathfrak{I}_1\mathfrak{I}_1+2\mathfrak{I}_0\mathfrak{I}_2) \\ \int d\nu\, g(\nu)\,(\mathfrak{I}_0\mathfrak{I}_0)(2\mathfrak{I}_0\mathfrak{I}_1) & \int d\nu\, g(\nu)\,(2\mathfrak{I}_0\mathfrak{I}_1)^2 & \int d\nu\, g(\nu)\,(2\mathfrak{I}_0\mathfrak{I}_1)(\mathfrak{I}_1\mathfrak{I}_1+2\mathfrak{I}_0\mathfrak{I}_2) \\ \int d\nu\, g(\nu)\,(\mathfrak{I}_0\mathfrak{I}_0)(\mathfrak{I}_1\mathfrak{I}_1+2\mathfrak{I}_0\mathfrak{I}_2) & \int d\nu\, g(\nu)\,(2\mathfrak{I}_0\mathfrak{I}_1)(\mathfrak{I}_1\mathfrak{I}_1+2\mathfrak{I}_0\mathfrak{I}_2) & \int d\nu\, g(\nu)\,(\mathfrak{I}_1\mathfrak{I}_1+2\mathfrak{I}_0\mathfrak{I}_2)^2 \end{pmatrix} \tag{24}$$

where $i, j = 0, 2, 4$ and

$$g(\nu) = \frac{1}{\text{Volume}} f(\nu) \left(\frac{M}{\bar{\rho}}\right)^3. \tag{25}$$

In this paper, we terminate this series after A_4 term. One can always go to higher order terms to get desired accuracy at higher k.

We will present the results of analytic calculations of A_{2n} in the next section. Calculating the variance of each of these coefficients is as straightforward as calculating the coefficient itself, performing the integrals over the halo mass function. We calculate the variance on these terms for a volume of 1 $(\text{Gpc}\,h^{-1})^3$ for different cosmological models (the 38 models explained in next section) at three different redshifts: 0.0, 0.5 and 1.0. Fig. 1 shows the relative variance of the three coefficients. We find for 1 $(\text{Gpc}\,h^{-1})^3$ the relative error σ_{A_0}/A_0 varies from 0.5 to 2 per cent, whereas on σ_{A_2}/A_2 and σ_{A_4}/A_4 vary from 1 to 7 per cent and from 2 to 20 per cent, respectively, depending on the cosmology and redshift. We see that the relative error on A_{2n} increases with n: this is a consequence of the fact that terms with higher n receiving a larger contribution from higher mass objects, since the mass scaling of the integrand for A_{2n} in the equations above is $M^{1+2n/3}$, while for the variance it is $M^{3+4n/3}$. Higher mass objects are rarer and their Poisson fluctuations are larger, hence the relative variance is increased. Below we will compute the sensitivity of these parameters to cosmology: we will show that A_0 contains most of the information on the amplitude σ_8. In this paper, we use halo mass function of Tinker et al. (2008).

So far we assumed $F(k) = 1$ without specifying its role. It was pointed out already in the original halo model (Seljak 2000) that the 1-halo term of the halo model fails to account for mass and momentum conservation at low k: the non-linear corrections to the power spectrum have to scale as k^4 or $-k^2P(k)$ at low k, while the leading order of the 1-halo term scales as k^0. At very low k such a term may even dominate over the linear term, which cannot be physical in the context of DM, even though it can happen in the context of galaxies Baldauf et al. (2013).

We will impose this constraint by simply fitting the residuals to the simulations at low k and apply the derived transfer function $F(k)$, which vanished at low k, to the model. We will show that the function $F(k)$ does not strongly depend on the cosmological model and we will thus ignore its dependence on cosmological parameters.

3 CALIBRATING THE MODEL WITH SIMULATIONS

We use cosmic emulator (Heitmann et al. 2009, 2010; Lawrence et al. 2010) to evaluate the power spectra for each of the 38 emulator simulations and assume in each case it gives the true non-linear matter power spectrum. These reference power spectra are correct to nearly 1 per cent up to $k \sim 1\,h\,\mathrm{Mpc}^{-1}$ at 38 different nodes (labelled as 0–37) in cosmological parameter space. This accuracy degrades to 5 per cent when computing the power spectrum away from the nodes. Node 0 cosmology is closest to the *Wilkinson Microwave Anisotropy Probe*-7 (*WMAP*-7) cosmology and we use it as a reference cosmology. We fit these simulation power spectra with our description – quasi-linear Zeldovich term plus modified 1-halo term as a sum of even powers of k, to determine coefficients A_{2n} as a function of cosmology.

To begin with, we fit the even power law (equation 14 with $F(k) = 1$) to the difference between matter power spectrum from emulator $P_{\rm Emu}$ and the Zeldovich term $P_{\rm Zel}$ for all 38 cosmologies and three redshifts: 0.0, 0.5, 1.0 between $k = 0.2$ and $0.8\,h\,\mathrm{Mpc}^{-1}$.

All the coefficients fitted, A_0, A_2 and A_4, are strongly correlated with σ_8, with A_0 having the least scatter. Fig. 2 shows the scaling of these coefficients with $\sigma_8(z)$ and $\sigma_{11.3}(z)$, where the latter was chosen to minimize the scatter in A_0. Each of these coefficients can be approximately fit as a power law irrespective of the redshift and cosmology, with $\sigma_8(z)$ scaling

$$A_0 \propto \sigma_8^{3.9},\ A_2 \propto \sigma_8^{3.0},\ A_4 \propto \sigma_8^{2.2}. \tag{26}$$

It is not straightforward to determine the errors since this is not a formal fit to a set of data points with individual errors. In Fig. 2, we also show results when the slope of A_0 is 4.0: we see this is also a good fit over the range.

Fig. 2 also shows the predictions of the halo model for these coefficients (in black crosses). While the halo mode predicts well A_0 at low redshifts, it fails for higher order coefficients. This can be improved if the virial radius is increased by roughly a factor of 2 at low redshifts, and more than that at higher redshifts (which needs to be taken to power $2n$ to evaluate the effect on A_{2n}), shown as red crosses in Fig. 2. The failure of the halo model to quantitatively predict these coefficients is not surprising: the haloes do not suddenly stop at the virial radius and the halo model has some flexibility in how it is implemented. Our goal here is not to understand the halo model, but to have accurate predictions. For this reason, we will just use the fits of A_{2n} coefficients to simulations in this paper.

The next step is to correct for the scatter around the best fit σ_8. A correlation is noticed between the residual of the coefficients with the effective slope $n_{\rm eff}$. This is shown in Fig. 3. Here, the residual means the difference between the diamond-bullets and best-fitting lines in Fig. 2 and effective slope $n_{\rm eff}$ is calculated as the slope of the linear matter power spectrum at $k \sim 0.2\,h\,\mathrm{Mpc}^{-1}$. The higher order coefficients have larger scatter and stronger correlation between this residual and effective slope. We tested the scalings for few different values of R in σ_R and found minimum scatter for $\sigma_{11.3}$, which can be seen in Figs 2 and 3. By using $\sigma_{11.3}$ instead of σ_8, one can remove the correlation with effective slope for A_0, so no $n_{\rm eff}$ correction is needed for A_0. However, A_2 and A_4 still need to be corrected for this correlation, although the correction is smaller in case of $\sigma_{11.3}$ than σ_8. Hence, the corrected expressions for these coefficients are

$$A_0 = 1529.87\sigma_8^{3.9} \times (1 + [-0.22 n_{\rm eff} - 0.4]),\ \text{ or } A_0 = 2167.39\sigma_{11.3}^{3.9}, \tag{27}$$

$$A_2 = 1299.75\sigma_8^{3.0} \times (1 + [-1.58 n_{\rm eff} - 2.8]),\ \text{ or } A_2 = 1724.16\sigma_{11.3}^{3.0} \times (1 + [-1.39 n_{\rm eff} - 2.5]), \tag{28}$$

$$A_4 = 758.31\sigma_8^{2.2} \times (1 + [-2.27 n_{\rm eff} - 4.2]),\ \text{ or } A_4 = 947.47\sigma_{11.3}^{2.3} \times (1 + [-2.12 n_{\rm eff} - 3.9]). \tag{29}$$

We still need to account for the mass conservation, which forces the 1-halo term to go to 0 at low k. In Fig. 4, we plot the ratio of the difference between $P_{\rm Emu}$ and $P_{\rm Zel}$ with $P_{\rm SimFit}$ which is given by $P_{\rm SimFit} = A_0 - A_2 k^2 + A_4 k^4$, where these coefficients are the best-fitting values to $(P_{\rm Emu} - P_{\rm Zel})$ for all 38 cosmologies (diamond bullets in Fig. 2). The top-left, top-right and bottom-left panel shows the same quantity at three different redshifts: 0.0, 0.5 and 1.0, respectively. All 38 curves in each panel are very close to 1 for k between 0.2 and $0.8\,h\,\mathrm{Mpc}^{-1}$, which is expected as these coefficients are fitted in that range in the first place. Outside this range the scatter increases. We took the average of all these 38 curves at all three redshifts and fit it to a 10th order polynomial, requiring to vanish at low k. The thick solid black line and dashed red curve represents the average and best fit to the average, respectively. It can be seen in bottom-right panel of Fig. 4 that these best fit to the average are very close to node 0 cosmology curve and also very close to each other for different redshifts for $k < 0.8\,h\,\mathrm{Mpc}^{-1}$. We average of these three best-fitting curves, at three different redshifts, to build the function $F(k)$ for the 1 halo term, which we model as

$$F(k) = \sum_{n=0}^{10} a_n k^n, \tag{30}$$

where the coefficients a_n are listed in Table 1. As expected by the mass conservation arguments, and seen in Fig. 4, this correction drops to zero for $k < 0.1\,h\,\mathrm{Mpc}^{-1}$. In principle we should force it to go to 0 as k^2, but we found this caused problems to the fit at higher k: the effects of $F(k)$ are very small in any case and in most instances below 1 per cent, since at low k the Zeldovich term dominates. For this reason, we will assume that this correction is independent of the cosmological model or redshift.

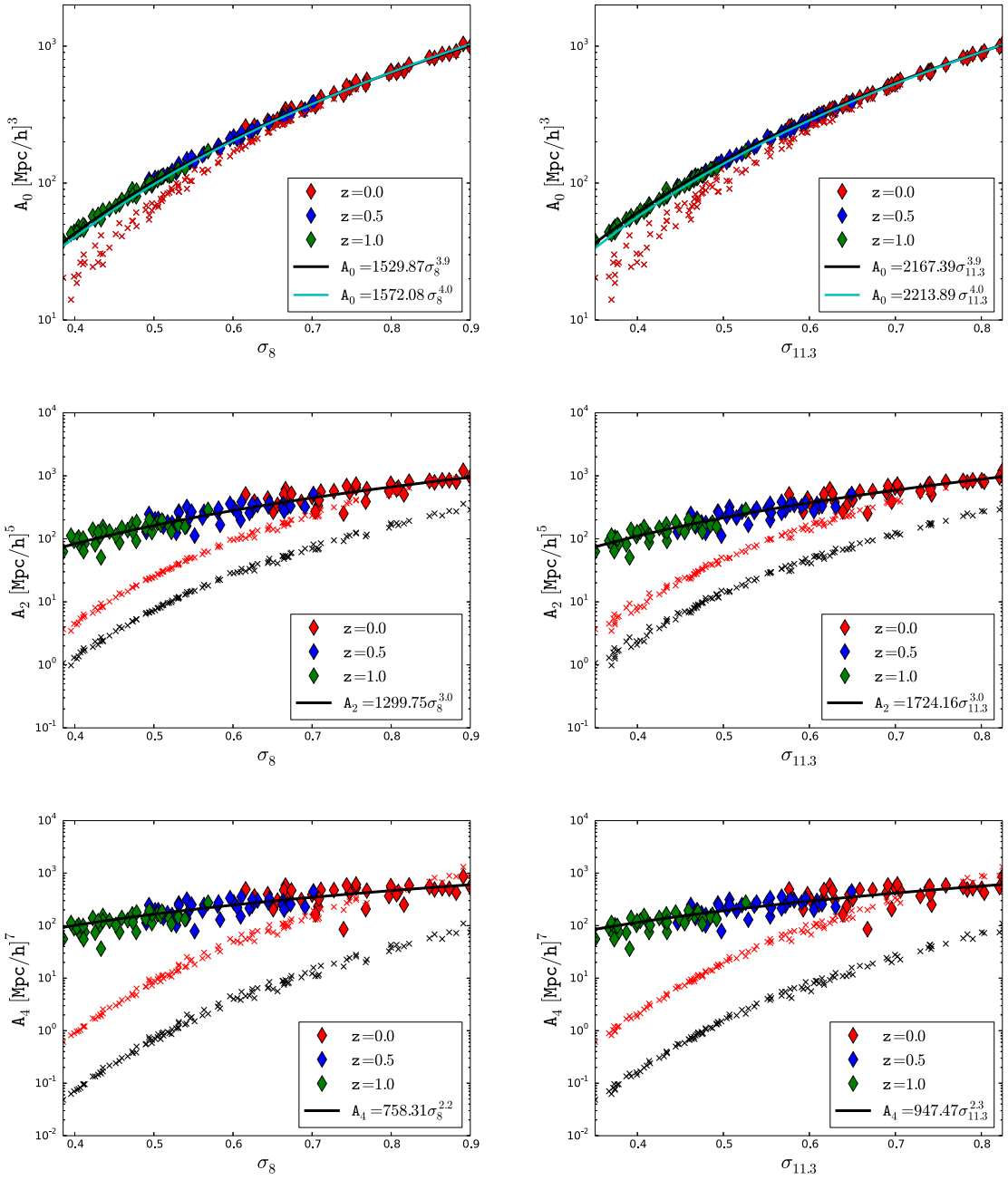

Figure 2. Fitted coefficients A_0, A_2 and A_4 versus σ_8 (left column) and $\sigma_{11.3}$ (right column). We see that $\sigma_{11.3}$ reduces the scatter relative to σ_8 for A_0. Solid black line is the best-fitting power law stated in the legend. The halo model prediction is shown in crosses, using the usual halo concentration parameter $R_s = R_{vir}/c$, with haloes extending to the virial radius R_{vir}, defined at the mean overdensity of 200 (black crosses), and doubling that to $2R_{vir}$ (red crosses). Halo model agrees well with simulations for A_0 at late redshifts, but not for A_2 and A_4 both in terms of amplitude and in terms of σ_8 or $\sigma_{11.3}$ scaling. Extending the halo profile to twice the virial radius improves the agreement.

We combine the above two terms to obtain the matter power spectrum as

$$P(k, z) = P_{\rm Zel}(k, z) + P_{\rm 1h}(k, z) \tag{31}$$

and,

$$P_{\rm 1h}(k, z) = (A_0 - A_2 k^2 + A_4 k^4) F(k), \tag{32}$$

where A_0, A_2 and A_4 are given by equations (27), (28) and (29), respectively, and $F(k)$ is given by equation (30).

We tested this expression against the matter power spectrum from emulator ($P_{\rm Emu}$) on 38 emulator nodes where the stated accuracy is 1 per cent. Fig. 5 shows the deviation of our predictions from the true matter power spectrum of Emulator at three different redshifts: 0.0, 0.5 and 1.0.

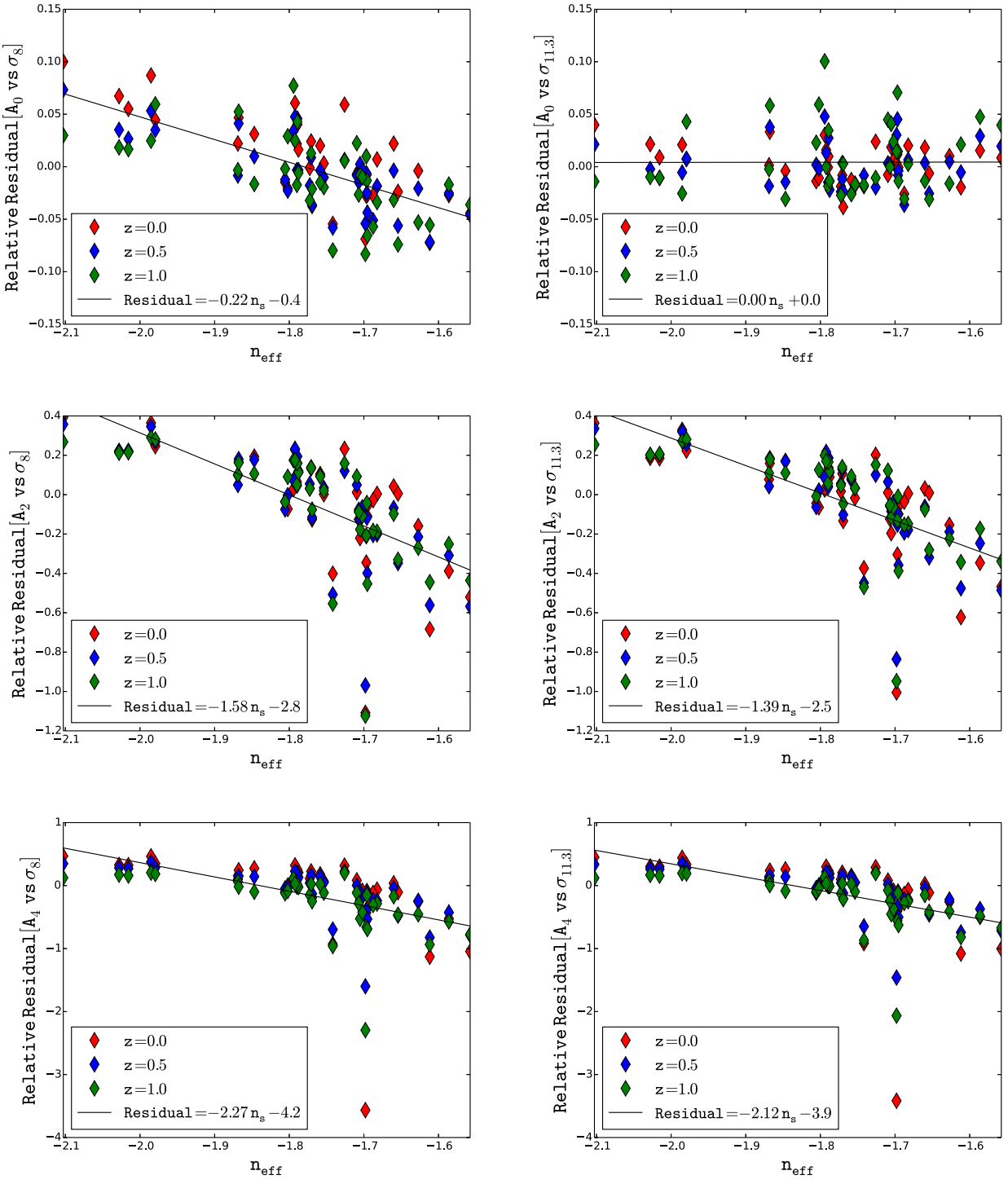

Figure 3. Correlation between effective slope ($n_{\rm eff}$) and residuals after σ_8 (left column) or $\sigma_{11.3}$ (right column) scaling is taken out and their respective best fit. Solid black line is the best linear fit as stated in the legend.

At redshift 0, we can predict the power spectra to a precision of 2–3 per cent up to $k \sim 0.5\,h\,{\rm Mpc}^{-1}$, except in some cosmologies which turn out to be unusual (typically equation of state very different from $w = -1$). At higher redshifts, this accuracy is even better for the same k, as expected since the non-linear effects are smaller. For most of the cosmological models, we can calculate these spectra to 5 per cent up to $k \sim 0.7\,h\,{\rm Mpc}^{-1}$ and much better for lower k.

In Fig. 6, we show the prediction of our model for *WMAP*-7 cosmology (node 0) with all its components plotted separately. Note that the $A_2k^2F(k)$ (in blue) term has a negative contribution while all other components have a positive contribution. The prediction of node 0 power spectrum is correct to about 2 per cent up to $k \sim 0.6\,h\,{\rm Mpc}^{-1}$ increasing to 4 per cent at $k \sim 0.7\,h\,{\rm Mpc}^{-1}$. This can also be seen in Fig. 5 where thick black line shows the ratio of the predicted and true matter power spectrum for node 0.

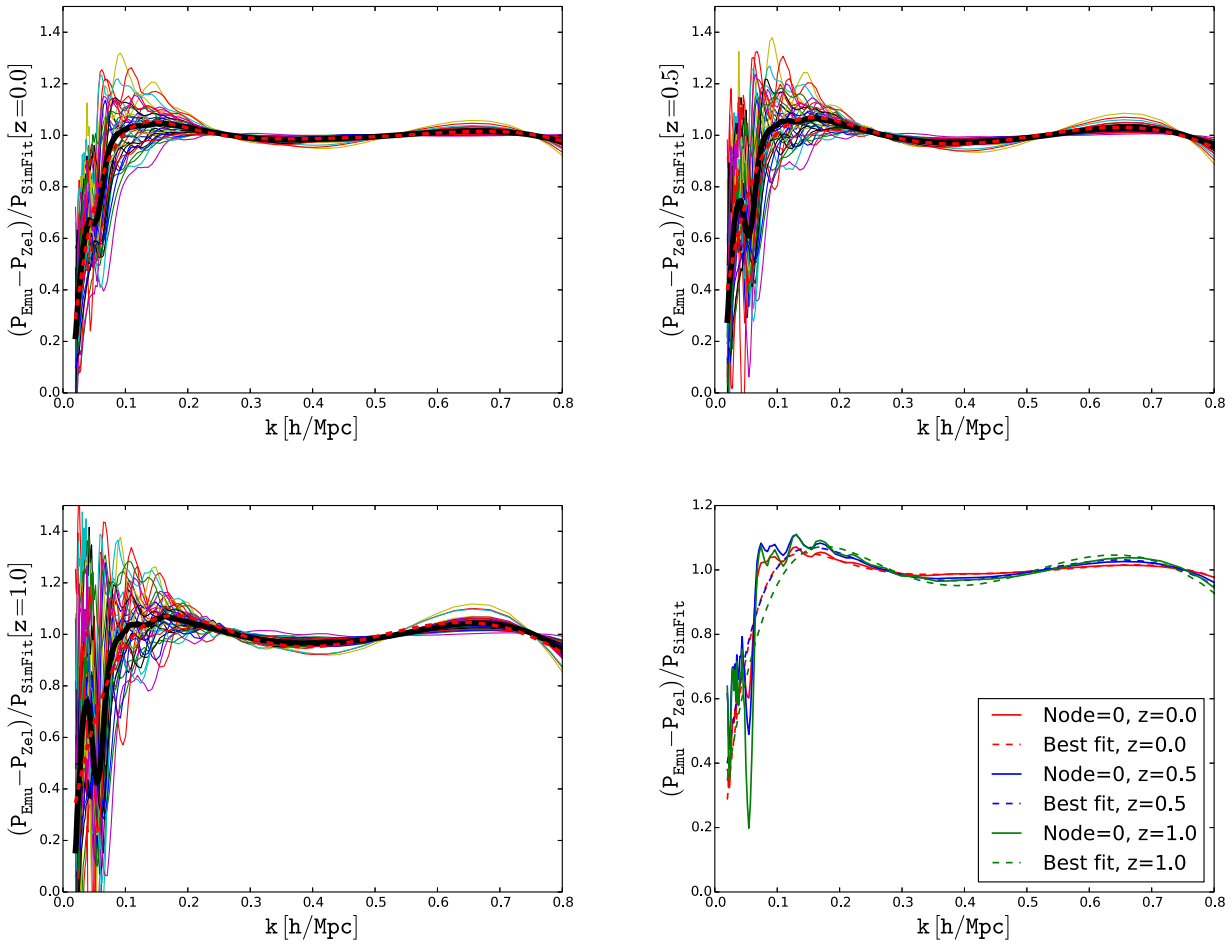

Figure 4. The first three panel (in reading order), shows the ratio of ($P_{\rm Emu} - P_{\rm Zel}$) and $P_{\rm SimFit} = A_0 - A_2k^2 + A_4k^4$, where the coefficients A_0, A_2, A_4 are the best-fitting coefficients to the emulator matter power spectrum for all 38 cosmological models (in different colours) at three different redshifts: 0.0 (top left), 0.5 (top right) and 1.0 (bottom left). Bottom-right panel shows same quantity for node 0 and the best fit to the average (of 38 coloured curves in first three panels) at three different redshifts.

Table 1. Coefficients to calculate the correction function, equation (30). The units of the coefficient a_n is (Mpc h^{-1})^n.

a_n	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
Value	0.0	21.814	−174.134	747.369	−2006.792	3588.808	−4316.241	3415.525	−1692.839	474.377	−57.228

We also explored how well can this expression predict the changes in the matter power spectrum when cosmological parameters are changed. We take emulator node 0 as the fiducial model and plot the relative difference with other nodes. The first three panel of Fig. 7 (in reading order) shows these derivatives for different components: linear term (in red), Zeldovich term (in green), emulator (in blue) and our predicted model (in thick black). Our predictions are matching very well with that of the true matter power spectrum from emulator, and certainly much better than pure linear theory or pure ZA. Note that we also get very good agreement of BAO smoothing, in contrast to linear theory predictions: this is because we are using ZA which smears out BAO. The broad-band effects of ZA are often anticorrelated with A_0: this is because an increase in σ_8 increases the non-linear smearing caused by the linear streaming of the displacement field, reducing the amplitude of the power spectrum in the ZA, while at the same time the amplitude of the A_0 is increased by the 1-halo term, generated by having more haloes at the same halo mass. The latter effect typically wins: the total power spectrum and the Zeldovich power spectrum are typically, but not always, on the opposite side relative to the linear power spectrum.

Of particular interest is the change in neutrino mass, also shown in Fig. 7. We compare the model predictions to the simulations of Bird, Viel & Haehnelt (2012). We see that our model predicts nearly perfectly the changes in the non-linear power spectrum induced by massive neutrinos. This shows that non-linear effects of massive neutrinos are no different than any other parameter: on large scales they follow linear theory, while on small scales the effects are dominated by the change in A_0. For $\sum m_\nu = 0.15$ eV the change in σ_8 is about 3 per cent and the corresponding change in $A_0 \propto \sigma_8^{3.9}$ is 13 per cent, while ZA goes in the opposite direction, so the linear suppression of 7 per cent at $k \sim 0.2\,h\,{\rm Mpc}^{-1}$ is increased to 11 per cent at $k \sim 0.8\,h\,{\rm Mpc}^{-1}$, in perfect agreement with simulations.

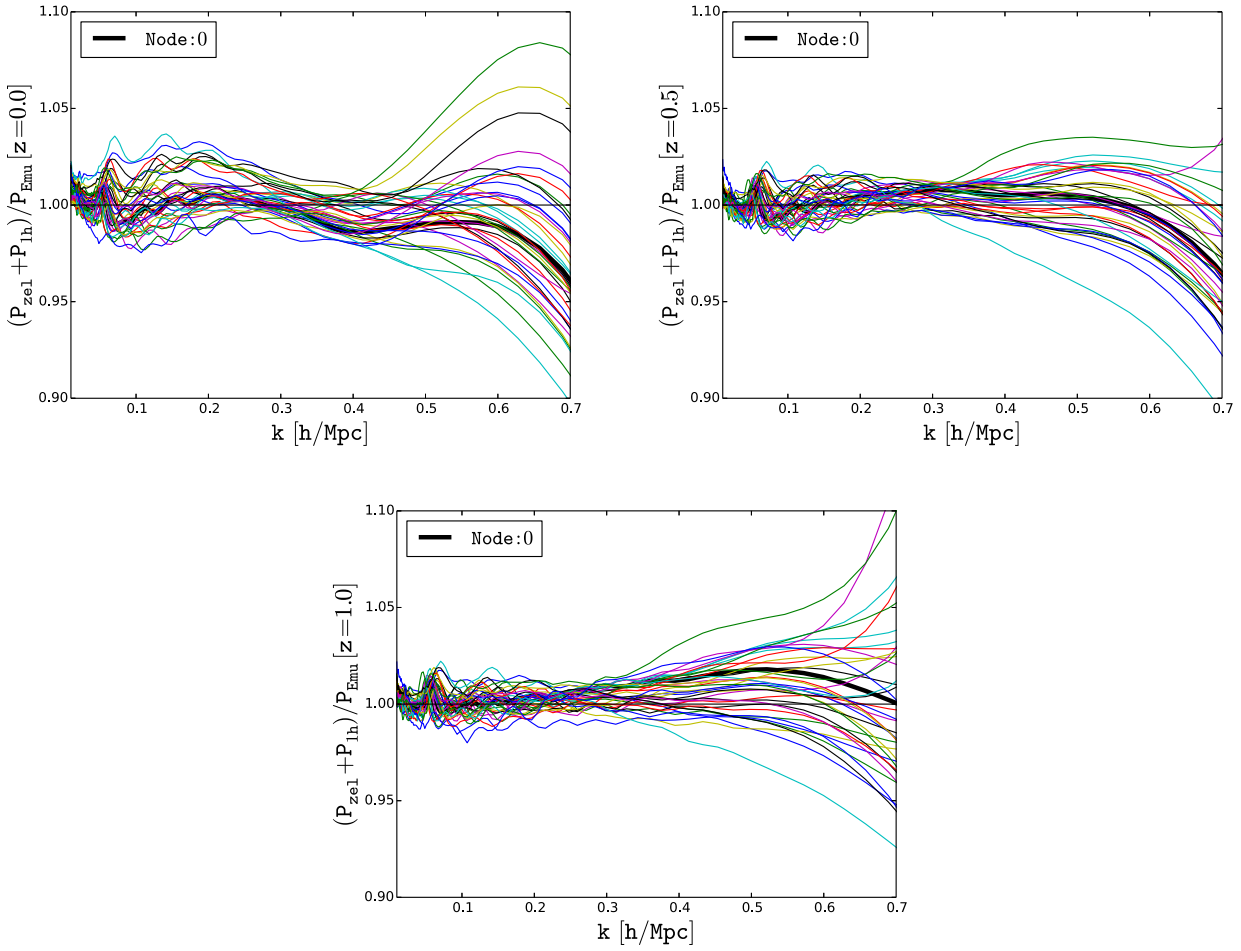

Figure 5. The residuals of our $P(k) = P_{\rm Zel}(k) + P_{\rm 1h}(k)$ expression against simulations for 38 different cosmological models (different colour curves in each panel) for three different redshifts.

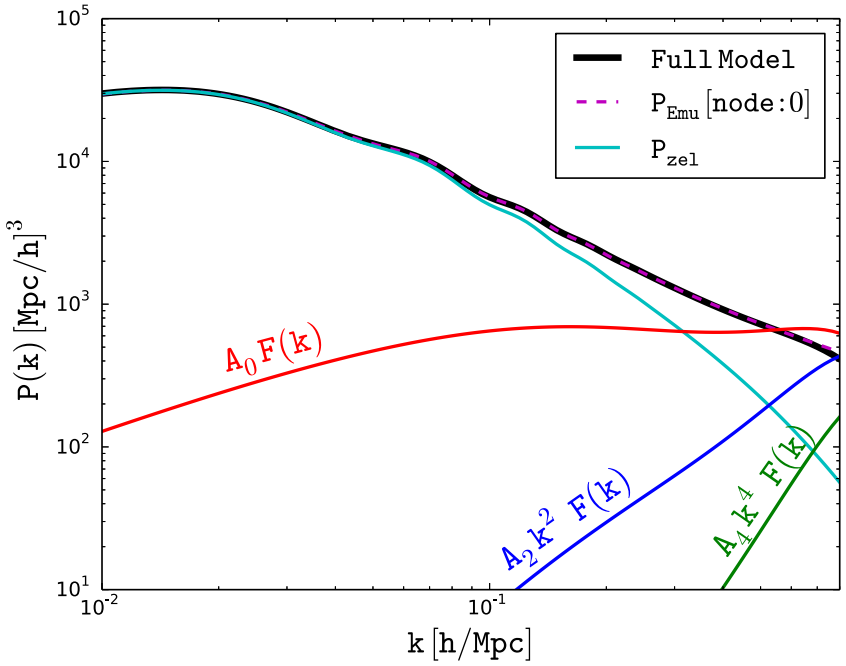

Figure 6. Matter power spectrum for *WMAP*-7 cosmology at redshift 0.0 from simulations (dashed magenta line) versus Zeldovich term (cyan line), $A_0F(k)$ term (red line), $A_2k^2F(k)$ term (blue line), $A_4k^4F(k)$ term (green line). Thick black line is the full predicted model from this work and is nearly indistinguishable from the simulations.

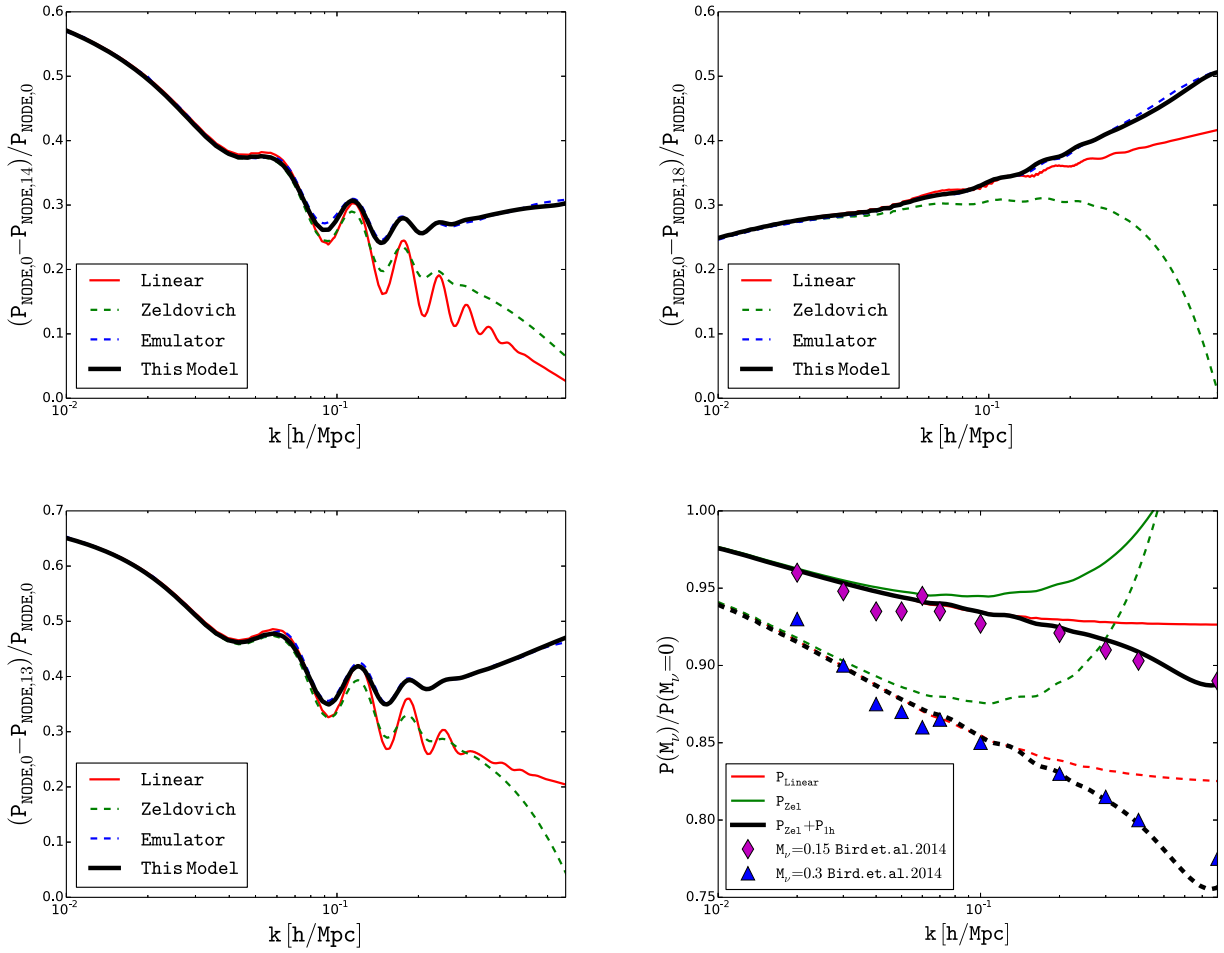

Figure 7. Relative difference in matter power spectrum between node 0 (Emulator) and node 14 (top-left), 18 (top-right), 13 (bottom-left). Showing the same quantity for linear term (in solid red), Zeldovich term (in dashed green), Emulator power spectrum (in dashed blue) and our prediction (in solid black). Bottom-right panel shows the ratio of the matter power spectrum with and without neutrino mass, for $\sum M_\nu$=0.15 (in solid lines) and 0.3 (in dashed lines) from Bird et al. (2012).

4 COVARIANCE MATRIX AND THE COSMOLOGICAL INFORMATION CONTENT OF $P(k)$

We next turn to the issue of covariance matrix. On large scales, low k, the covariance matrix is based on Gaussian approximation. As we move to higher k, the modes become correlated and the covariance matrix becomes non-Gaussian. In our model, the non-Gaussianity comes from two separate terms. First is the non-Gaussian nature of the Zeldovich term and second is the non-Gaussian nature of the 1-halo term. We will not analyse the non-Gaussian covariance matrix in ZA in this paper, as there are currently no analytic calculations available. We also do not have any analytic predictions for the correlation between the Zeldovich part and the 1-halo part. For the 1-halo term, we will focus on A_0 contribution, since as we will argue in next section we should marginalize over the higher order terms anyways. In our initial discussion, we will ignore the supersample variance contribution (Takada & Hu 2013; Li et al. 2014), which will be discussed separately below.

The halo model calculations in Fig. 1 suggest that the relative variance σ_{A_0}/A_0 should be around $0.01\sqrt{(\mathrm{Gpc}\,h^{-1})^3/\mathrm{Volume}}$, depending on the cosmological model and redshift. This calculation is given by

$$\left(\frac{\sigma_{A_0}}{A_0}\right)^2 = \frac{\int f(\nu)\mathrm{d}\nu M^3}{[\int f(\nu)\mathrm{d}\nu M]^2 \bar{\rho} V}, \tag{33}$$

and is determined by the fourth moment of mass integrated over the halo mass function and thus very sensitive to the halo mass function accuracy at the high-mass end. Just as in the case of the halo model predictions for the scalings of A_0, A_2 and A_4, we may not completely trust the halo model predictions. We will write the following ansatz to the covariance matrix $\mathrm{Cov}(P(k_i), P(k_j)) = \langle P(k_i)P(k_j)\rangle - \langle P(k_i)\rangle\langle P(k_j)\rangle$,

$$\mathbf{Cov}(P(k_i), P(k_j)) = P(k_i)P(k_j)\left(\frac{2}{N_i}\delta_{ij} + \left(\frac{\sigma_{A_0}}{A_0}\right)^2\right). \tag{34}$$

Here, N_i is the number of Fourier modes in the ith bin. Our model predicts that the scaling of the variance is

$$\frac{\sigma_{A_0}}{A_0} = \frac{\delta_{A_0}}{[(V/1h^{-1}\mathrm{Gpc})^3]^{1/2}},\ \delta_{A_0} = 0.0079(h^{-1}\mathrm{Gpc})^{3/2}, \tag{35}$$

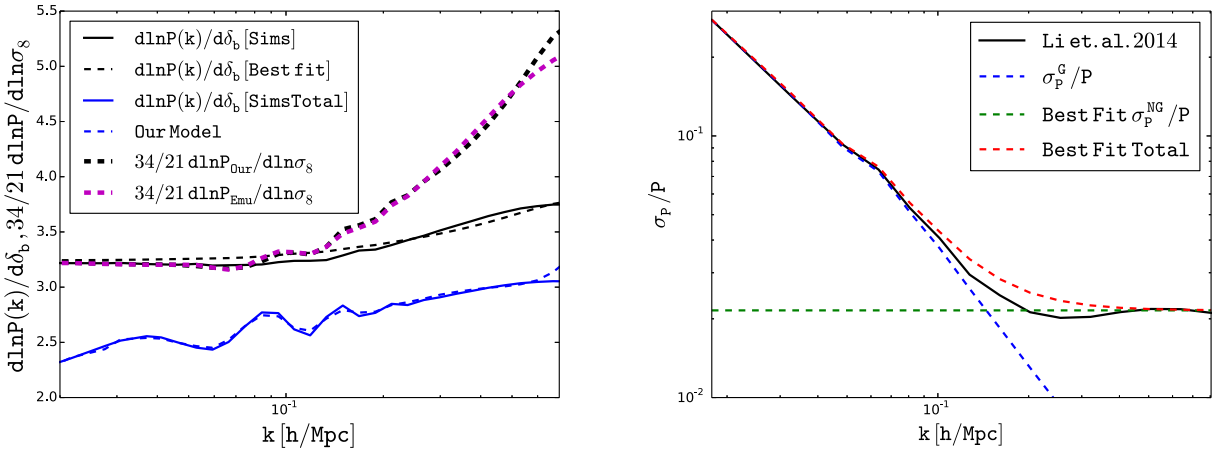

Figure 8. Left: derivative of the matter power spectrum with respect to the change in curvature (i.e. background density) from simulations (blue solid line) and our best-fitting model (blue dashed line). The same, but only for the growth effect without dilation, is shown with the corresponding black solid and black dashed lines. The thick dashed lines show the derivative with respect to the amplitude change, such that it is degenerate with the curvature change at low k: shown are the predictions from simulations (in magenta) and from our model (in black). We see that the degeneracy is broken at higher k even in the absence of the dilation effect. Right: relative variance in the matter power spectrum: $\sqrt{2/N}$ (blue dashed line) where N is the number of modes, best fit σ_p/P (green dashed line), and the total (red dashed line) as the norm of the two terms.

where V is the volume in units of $(h^{-1}\mathrm{Gpc})^3$, and the value of $\delta_{A_0} = 0.0079(h^{-1}\mathrm{Gpc})^{3/2}$ was obtained from a fit of the model to the diagonal part of the covariance matrix derived from *Planck* cosmology simulations in Li et al. (2014), shown in the right-hand panel of Fig. 8. This value is slightly lower than the predictions of the halo model in Fig. 1. Since the predictions are very sensitive to the massive end of the halo mass function, which is not well determined, we should not expect perfect agreement.

It is important to note that the covariance matrix depends on the simulated volume: if the volume changes the covariance matrix will change, and this means that comparing one set of covariance matrix results to another is not trivial. We can simplify the expression if we express the number of modes in terms of a fixed width of the k bin Δk, $N = 4\pi k^2 \Delta k V/(2\pi)^3$. One can see that both the Gaussian sampling variance term and the Poisson term scale with volume, so that

$$\mathbf{Cov}(P(k_i), P(k_j)) = P(k_i)P(k_j)V^{-1}\left(\frac{4\pi^2}{k_i^2\Delta k}\delta_{ij} + \delta_{A_0}^2\right). \tag{36}$$

The relative contribution of diagonal versus off-diagonal terms still depends on the width of the binning in k, but the overall volume scaling is the same.

Now that we have fixed the only free parameter of our model δ_{A_0}, we can apply it to another set of simulations to see the agreement. We have compared it to results in Blot et al. (2014), which used 12 288 boxes of size 656.25 h^{-1} Mpc to derive the full covariance matrix. In Fig. 9 (upper panels), we have compared our model to these simulations for both diagonal and off-diagonal parts of the covariance matrix. We show that the diagonal part of the covariance matrix (left-hand panel of Fig. 9) is an excellent fit, even better than comparison with Li et al. (2014), and this is without any free parameters. In the right-hand panel of Fig. 9, we show the off diagonal terms for six different k values. Our model predicts the off-diagonal correlation coefficients are simply a constant, except at the diagonal where there is an additional Gaussian contribution. Our prediction is in reasonable agreement with these simulations: we are able to reproduce simulation results for both diagonal and off-diagonal terms to within 10–20 per cent, which is remarkable given its simple form and no free parameters.

4.1 Variance of the covariance matrix

An interesting and important question is how big do the simulations need to be to converge. For the convergence of the power spectrum, the answer is given by $\sigma_{A_0}/A_0 = \delta_{A_0}/V^{1/2}$ and we can see that $V = 1\ (h^{-1}\ \mathrm{Gpc})^3$ is sufficient for 1 per cent accuracy. For the covariance matrix, this requirement becomes considerably stricter. One can write an expression for the relative variance of the covariance term as

$$\left(\frac{\sigma(\sigma_{A_0})}{\sigma_{A_0}}\right)^2 = \frac{\int f(\nu)\mathrm{d}\nu M^7}{[(\int f(\nu)\mathrm{d}\nu M^3]^2 \bar{\rho} V}, \tag{37}$$

so we can see that this is given by the eighth moment of the mass averaged over the halo mass function. The results of this prediction are shown in Fig. 10. The rms variance for $V = 1\ (h^{-1}\ \mathrm{Gpc})^3$ is now about 10–30 per cent and the corresponding error on the covariance matrix (which goes as a square of σ_{A_0}) is thus 20–70 per cent. There is a large spread in the value because the calculation is so sensitive to the very high mass end of the halo mass function, which is poorly known, so the resulting values should only be taken as indicative and can probably vary by a factor of 2. This is simple to understand: occasionally there will be a large cluster formed which will significantly change the value of A_0, and consequently make its variance change considerably.

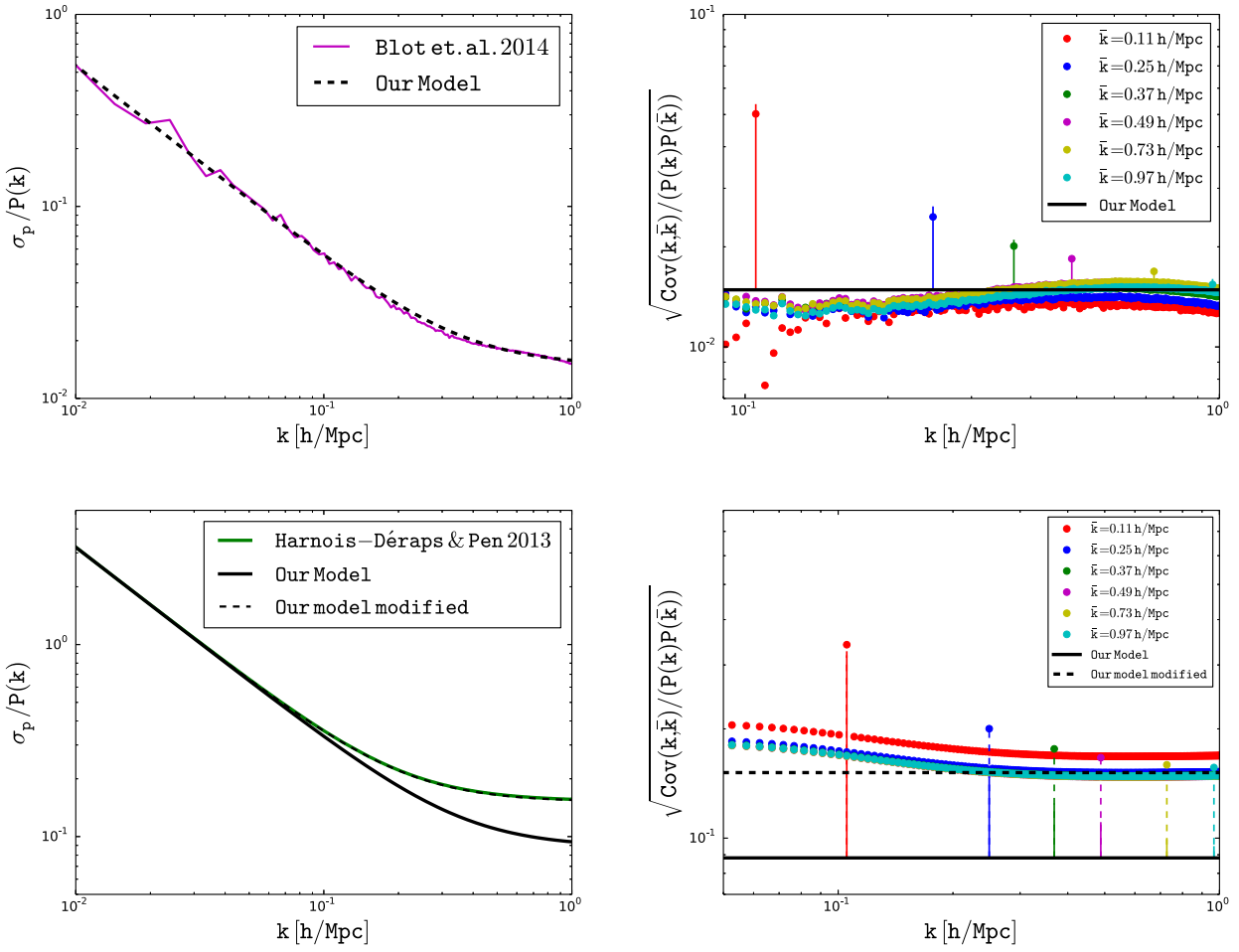

Figure 9. Comparison between our model prediction of covariance matrix with Blot et al. 2014 (upper panels) and Harnois-Déraps & Pen (2012, lower panels) for diagonal (left-hand panels) and off-diagonal elements (right-hand panels). Note that there are no free parameters in the top, while for the bottom panel we show both our best model without a free parameter as well as a modified model where we fit for the value of σ_{A_0}/A_0, which is a valid procedure for these simulations, as discussed in the text. Our covariance matrix model (equation 34) is very simple, yet it is able to reproduce the full covariance matrix from simulations to within 10–20 per cent.

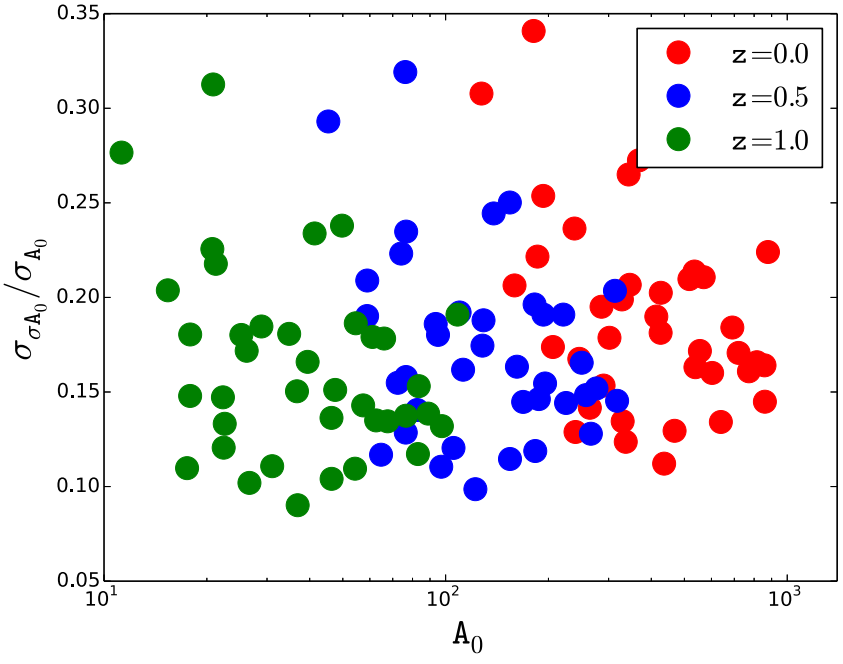

Figure 10. Relative variance $\sigma(\sigma_{A_0})/\sigma_{A_0}$ versus A_0 based on our model for three different redshifts: 0.0 (red), 0.5 (blue) and 1.0 (green). Each bullet is one cosmological realization of the 38 cosmic emulator nodes.

As an example, when we compare our model predictions of the covariance matrix to Harnois-Déraps & Pen (2012), we find that the agreement is not very good, in that our model predicts lower covariance matrix than measured, and the predicted value of σ_{A_0}/A_0 is about 40 per cent below the required for fit the simulations. However, Harnois-Déraps & Pen (2012) used a total simulated volume of 1.6 $(h^{-1}\,\mathrm{Gpc})^3$, suggesting that the value of σ_{A_0}/A_0 has only been determined to about 10–25 per cent. If we let the value of σ_{A_0}/A_0 to be free, we again find a remarkable agreement with the simulations.

To converge on the covariance matrix at 1 per cent one needs a simulated volume to be of the order of 500–5000 $(h^{-1}\,\mathrm{Gpc})^3$. This is an enormous volume: it explains why in recent work of Blot et al. (2014) they needed to simulate 12 288 simulations with a total volume of 3350 $(h^{-1}\,\mathrm{Gpc})^3$ to converge.

4.2 Information content

We can now combine the variance of A_0 with its scaling with σ_8, $A_0 \propto \sigma_8^{3.9}$, to derive the cosmology information content of the A_0 term,

$$\frac{\sigma_{\sigma_8}}{\sigma_8} = \frac{\sigma_{A_0}}{3.9A_0} = 0.002\sqrt{(h^{-1}\,\mathrm{Gpc})^3/\mathrm{Volume}}. \tag{38}$$

This is a remarkably small number, which suggests that much of the cosmological information on the rate of growth of structure, and consequently on the Figure of Merit for dark energy equation of state (Mortonson, Huterer & Hu 2010), resides in this term. To achieve a comparable precision on linear scales one would need about 5×10^5 modes, which for 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume would correspond to $k_{\mathrm{max}} = 0.31\,h\,\mathrm{Mpc}^{-1}$. This is already well into the non-linear regime for $z < 1$ implying that we do not have this number of linear modes available, so the bulk of the cosmological information on the amplitude comes from A_0 term. However, since A_0 is mostly sensitive to amplitude (best correlation is with $\sigma_{11.3}$) and nothing else, this also suggests that information on other parameters that depend on the shape of $P(k)$ and not its amplitude will be less well determined.

While we do not have reliable variance predictions for A_2 and A_4 from simulations, Fig. 1 suggests that A_2 has variance three times larger than A_0 and A_4 has variance another three times larger than A_2. This is mostly caused by the fact that Poisson fluctuations get larger for higher order coefficients because of their mass weighting: for example, A_2 weighting is $M^{5/3}$ as opposed to M for A_0, giving more weight to higher mass haloes, which are rarer and therefore have larger Poisson fluctuations. This combined with less steep scaling of A_2 and A_4 with σ_8 compared to A_0 (equations 27, 28 and 29), suggests that there is little additional information in these two coefficients. Another argument for why information in A_2 and A_4 should be ignored, based on baryonic effects, will be presented below.

4.3 Supersample variance

Supersample variance (Hamilton, Rimes & Scoccimarro 2006) arises from the very long wavelength density modes that appear as constant on the scale of the survey. These can be viewed as a change of curvature inside the observed volume (Baldauf et al. 2011), and this couples to all the short wavelength modes. On large scales the effect can mimic a change in the amplitude of fluctuations, together with a rescaling of the length (Sherwin & Zaldarriaga 2012):

$$\delta \ln P(k) = \left(\frac{47}{21} - \frac{1}{3}\frac{\mathrm{d}\ln P}{\mathrm{d}\ln k}\right)\delta_b = \left(\frac{68}{21} - \frac{1}{3}\frac{\mathrm{d}\ln(k^3P)}{\mathrm{d}\ln k}\right)\delta_b, \tag{39}$$

where δ_b is the density perturbation on the scale of the survey volume. The first term is the effect of the curvature on the growth of small-scale modes, while the second term is the dilation due to the presence of local curvature. It is important to recognize that on large scales the growth effect is degenerate with a $(34/21)\delta_b$ change of amplitude σ_8, while the dilation effect of $-\delta_b/3$ is degenerate with a change in scale, i.e. with a change in the angular diameter distance that can arise from a change in cosmological parameters. We will assume that the change in scale cannot be used as an indicator of the supersample variance because of its degeneracy with these other parameters, so we will only focus on the change in growth rate. The rms fluctuations of $1(\mathrm{Gpc}\,h^{-1})^3$ volume are about 0.4 per cent (Takada & Hu 2013), which together with the 34/21 factor implies that at low k one cannot determine σ_8 to better than 0.6 per cent in the linear regime, which is a factor of 3 larger error than the error on σ_8 without the supersample variance in equation (38). It is therefore clear that without addressing this issue the supersample variance dominates the errors.

On smaller scales we expect the non-linear effects are no longer degenerate with a change in σ_8. Physically the reason for difference is in the curvature nature of the supersample variance: curvature effects grow with the growth rate, i.e. the growth of short wavelength mode δ_s due to the coupling to the long wavelength mode scales as $\delta_s(z)[1 + 34D(z)\delta_{b0}/21]$, where $D(z)$ is the linear growth rate and δ_{b0} is the long wavelength mode today, and thus this coupling only matters at low redshifts since $D(z) \ll 1$ for $z \gg 1$. This is different from a simple change in overall amplitude $\delta_s(z)(1 + \delta\sigma_8)$, which has no redshift dependence.

To understand this more quantitatively we can compute the logarithmic derivative of A_0 (equation 23) with respect to the two parameters in the context of the universal halo mass function $f(\nu)$, where ν is given by equation 6 (Slosar et al. 2008). The Lagrangian bias is defined as $b_{\mathrm{L}} = \bar{n}^{-1}\partial n/\partial\delta_b$, which can be rewritten using $\nu = (\delta_c - \delta_b)^2/\sigma^2$ as $b_{\mathrm{L}} = (-2\nu/\delta_c)\partial\ln[\nu f(\nu)]/\partial\nu$. In addition we also have the mean

density increased by δ_b inside the patch. We are still dividing the density with the global mean density, so $\bar{\rho}$ does not change. Using this we find

$$\frac{\mathrm{d}\ln A_0}{\mathrm{d}\delta_b} = \frac{\int (1+b_{\rm L}(\nu))\nu f(\nu) M \mathrm{d}\ln\nu}{\int \nu f(\nu) M \mathrm{d}\ln\nu} = \langle (1+b_{\rm L})\rangle. \tag{40}$$

So the logarithmic slope of A_0 with respect to a long wavelength modulation is given by the appropriate average of the Eulerian bias $b_{\rm E} = 1 + b_{\rm L}$.

If instead one looks at the logarithmic growth of the amplitude with respect to amplitude $\sigma_8 \propto \sigma(M)$, then $\mathrm{d}\nu/\mathrm{dln}\,\sigma_8 = -2\nu$, and so

$$\frac{\mathrm{d}\ln A_0}{\mathrm{d}\ln\sigma_8} = \delta_c \frac{\int b_L(\nu)\nu f(\nu) M \mathrm{d}\ln\nu}{\int \nu f(\nu) M \mathrm{d}\ln\nu} = \delta_c \langle b_{\rm L}\rangle. \tag{41}$$

Since $\mathrm{dln}\,A_0/\mathrm{dln}\,\sigma_8 = 3.9$ we find $\mathrm{dln}\,A_0/\mathrm{d}\delta_b = 3.9/1.68 + 1 = 3.3$. The response to the long wavelength mode has thus a lower logarithmic slope of growth relative to σ_8 and is not much larger than the linear regime value 68/21. This should be contrasted against the response to the amplitude change, which goes from σ_8^2 in the linear regime to $\sigma_8^{3.9}$ in the non-linear regime. Note that this calculation is valid if the density is divided by the global background density, as appropriate for weak lensing observations, which are sensitive to the total density. Whenever the density perturbation is defined using local mean density these numbers should be reduced by 2.

Numerical results are shown in the left-hand panel of Fig. 8, where we show the non-linear response to δ_b from simulations of Li et al. (2014), and the corresponding response to a change in σ_8 that mimics δ_b at low k. We can still model a change in δ_b as a quasi-linear term and $(A_0 - A_2k^2 + A_4k^4)F(k)$. For the quasi-linear term, we adopt simply the ZA model multiplied with the corresponding linear factor of $68/21\delta_b$, and we fit for the other three parameters. The result is shown in Fig. 8 and provides a reasonable fit to the simulations. Note that we show results with and without $\mathrm{dln}\,P/\mathrm{dln}\,k$ term, against simulations with and without it (Li et al. 2014). We find that for $\delta_b = 0.02$ A_0 has changed by 7.4 per cent, while the quasi-linear term has changed by 6.4 per cent, so that $\mathrm{dln}\,P/\mathrm{dln}\,\delta_b = 3.2$ at low k and 3.7 around $k \sim 0.5\,h\,\mathrm{Mpc}^{-1}$ where A_0 dominates. This is in reasonable agreement with the analytic estimate of 3.3. For the σ_8 scaling a change of 6.4 per cent in the linear term corresponds to 13 per cent change in A_0. The contrast between the two effects is shown in Fig. 8. The supersample variance is thus not degenerate with σ_8, so if one can determine both the quasi-linear term and A_0 term with sufficient accuracy, one can break the degeneracy between the two effects.

How well can one determine σ_8 in the presence of supersample variance? If we only have information from A_0, then the analysis above suggests that one can determine σ_8 to about $(3.7/3.9)0.4 \sim 0.38$ per cent in 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume, about a factor of 2 worse than without the supersample variance. If we have information both from linear regime and from A_0 dominated regime, then we can break the degeneracy between the supersample variance and σ_8. The extent to which this can be achieved depends on how well we can measure the amplitude in the linear regime: to reach 0.4 per cent accuracy we would need to measure all the modes up to $k \sim 0.2\,h\,\mathrm{Mpc}^{-1}$ in a 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume, which seems possible to achieve. Moreover, we note that a change in curvature cannot be modelled well with just a change in linear term and A_0, higher order terms also change significantly. Even though we argue below that these effects are degenerate with baryonic effects, this degeneracy may be broken in this situation given how different these effects are and given that there is a lot of information present at high k. In summary, the amplitude of fluctuations in a 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume can be determined to an accuracy of 0.4 per cent if the supersample variance cannot be determined, which can be reduced by a factor of 2 if the degeneracy between the supersample variance and σ_8 amplitude can be broken.

Instead of including the supersample variance effect in the covariance matrix, one can include it as an additional curvature parameter that one can marginalize over. The parameter is δ_b and its prior should be a Gaussian with a zero mean and rms variance σ_V determined by the survey window (see Takada & Hu 2013; Takada & Spergel 2014 for predictions for simple survey geometries). The response of the power spectrum to the long wavelength δ_b parameter should be

$$\delta P = \left(\frac{47}{21}P_{\rm Zel} - \frac{1}{3}\frac{\mathrm{d}P}{\mathrm{d}\ln k} + \left[3.7A_0 - 3A_2k^2 + 2.5A_4k^4\right]F(k)\right)\delta_b, \tag{42}$$

where $P_{\rm Zel}$, A_0, A_2 and A_4 are the values of the fiducial model around which we are exploring the supersample variance effect. For example, in a Markov-Chain Monte-Carlo (MCMC) chain this would be the model one is testing at a given chain position. We found that the fit to the simulations must include A_2 and A_4 terms and that the fit is only valid to $k \sim 0.7\,h\,\mathrm{Mpc}^{-1}$. Note that the change of A_2 and A_4 relative to A_0 is similar to that of amplitude change in equation (26).

5 EFFECTS OF BARYONS

Baryonic effects inside the DM haloes change the matter power spectrum relative to the DM alone and these effects must be incorporated into the analysis, otherwise they can lead to substantial bias in the cosmological parameter estimation (Semboloni et al. 2011, 2013). Baryonic effects can come in different forms. First is simply the fact that gas distribution inside DM haloes is distributed differently than the DM, because gas is hot and has significant pressure. As a result, gas has a core at the centre of the cluster, leading to reduced clustering strength on small scales. Second effect is baryon cooling, which causes gas to cool and condense into galaxies at the DM halo centres. This leads to an enhancement of the clustering relative to pure DM case. Baryons can also be pushed out of the halo centres by processes such as supernova and AGN feedback, which can in some cases push the gas quite far out. Furthermore, in all of these examples DM may also be redistributed

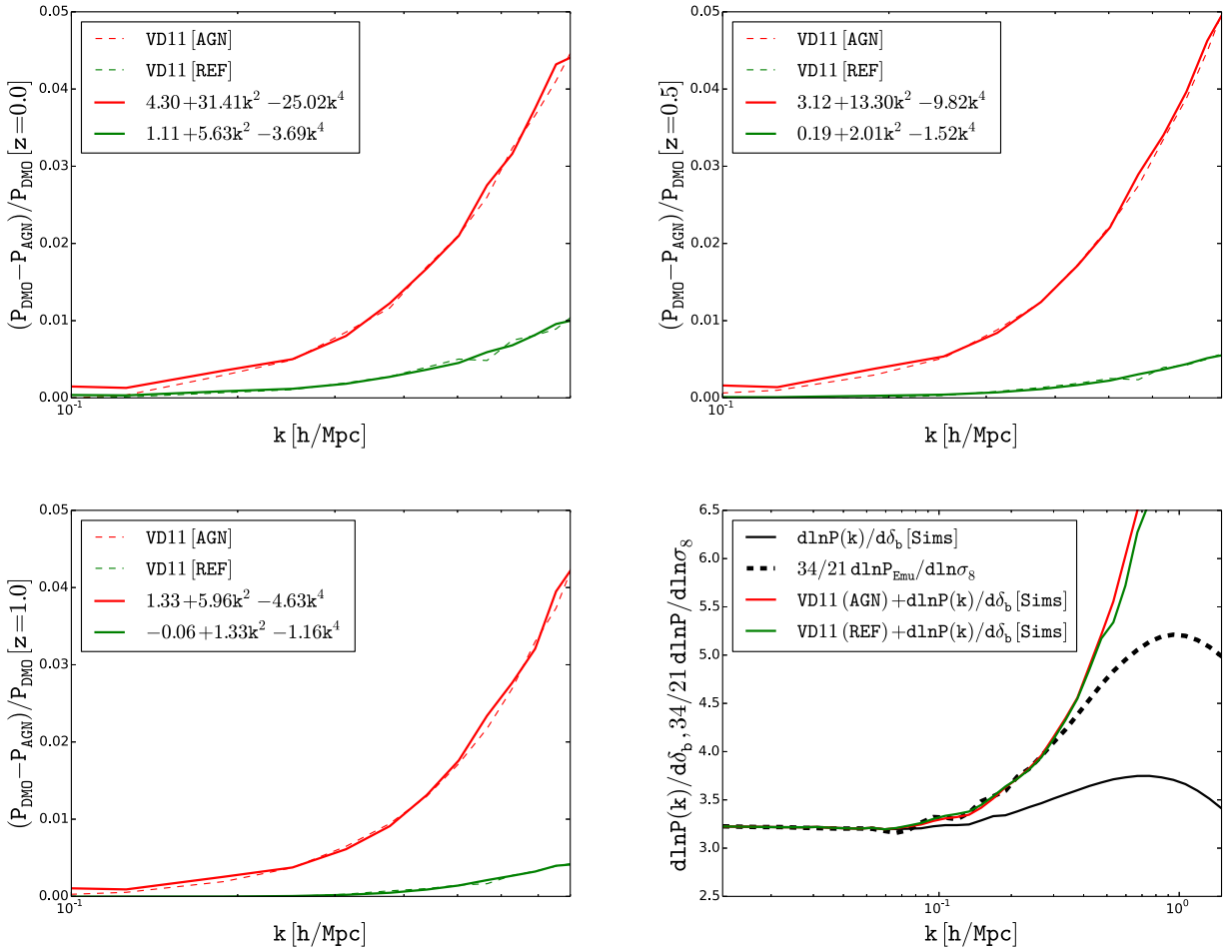

Figure 11. The first three panel (in reading order) are the relative difference between DMO model and AGN (red dashed line) and REF (green dashed line) from van Daalen et al. 2011 (VD11) at redshift 0.0 (top-left), 0.5 (top right) and 1.0 (bottom left). REF model contains the baryonic physics without any AGN feedback model. Solid lines (red and green) are the corresponding best fit $\delta A_0 - \delta A_2 k^2 + \delta A_4 k^4$ as explained in Section 5. Bottom-right panel shows the derivative of the matter power spectrum with respect to a change in background density (2 per cent) in solid-black from Li et al. 2014, and with respect to a change in amplitude using prediction from emulator (thick dashed black line).

as a consequence of the baryons either condensing on to the halo centres or being pushed out. For example, for baryonic cooling on to a galactic disc this process is known as adiabatic contraction (Blumenthal et al. 1984).

From the halo model point of view, the main effect of the baryons is the redistribution of the gas, and possibly DM, inside the haloes. This can be qualitatively described as the change in the scale radius R_s. The total mass of the halo M is unchanged, since these baryonic processes do not push the gas or the DM far out of the virial radius of the halo such that the halo mass would be affected. As a consequence, we expect that A_0 parameter is essentially unchanged, while A_2, A_4 etc. will change during the baryonic redistribution of matter.

To investigate this further we used simulation based matter power spectra from van Daalen et al. (2011) to compute the effects of baryons on the coefficients A_0, A_2 and A_4. In particular, we use the DM only and the supernova and AGN feedback models, corresponding to hydrodynamical simulations with supernova or AGN feedback model. It was argued that the latter is needed to reproduce cluster observations such as X-ray luminosity–temperature relation (McCarthy et al. 2010). We use the AGN model as the main model since it provides the largest effects, but we also explore reference supernova feedback model from van Daalen et al. (2011). Baryon corrections to the matter power spectrum from AGN feedback model exceed 1 per cent level for $k > 0.3\,h\,\mathrm{Mpc}^{-1}$ (van Daalen et al. 2011). We use the results at three different redshifts: 0.0, 0.5 and 1.0. In Fig. 11, we try to fit the difference between the pure DM and AGN model, $P_{\rm DMO} - P_{\rm AGN}$, or reference supernova feedback model $P_{\rm DMO} - P_{\rm REF}$, with the model $\delta A_0 + \delta A_2 k^2 + \delta A_4 k^4$, to estimate the changes in these coefficients due to baryons at each redshift (note that since the changes are only important at high k we can set $F(k) = 1$). We fit these models over the k range between 0.2 and $0.8\,h\,\mathrm{Mpc}^{-1}$. Fig. 11 shows the best-fitting models, which are a good fit to the simulations over this range. We also calculated these coefficients for the cosmology assumed in this paper using the results from Fig. 2.

We find that for the AGN model the relative change in A_0 is about 0.5–1 per cent, depending on the redshift, whereas the changes in A_2 and A_4 are about 4–7 and 4–8 per cent, respectively. If we assume no change in A_0 the fit is a bit worse and the change in A_2 and A_4 is larger. This confirms that the coefficient A_0 is quite indifferent to baryonic effects, while A_2 and A_4 are significantly more contaminated. The change is positive. This is expected since AGN feedback expands the gas and makes the scale radius R_s larger. It is less obvious why A_0 should increase

when gas is being pushed outwards, but the effect on A_0 is small and it could also be driven by the numerical fitting procedure. If we assume that the baryonic uncertainty is at the level suggested by these AGN models, then using equation (24) the corresponding uncertainty on σ_8 will be 0.5–1 per cent/3.9 $\sim$ 0.1–0.2 per cent from A_0, and about an order of magnitude larger from A_2 and A_4. Given that the difference between AGN and DM models is probably an overestimate of the error associated with the baryonic effects our analysis suggests that these effects can be effectively marginalized over without any loss of cosmological information from A_0. We also note that other baryonic feedback models from van Daalen et al. (2011), such as the reference model, while giving a lower amplitude of the effect, have very similar k-dependence, as can be seen from Fig. 11.

Above we argued that supersample variance effect should not be treated as a variance but as a separate parameter that can be determined from the data. Using linear theory and A_0 may not contain enough information to break the degeneracy between the amplitude σ_8 and supersample variance. Using higher k information may be more promising, since the two effects also have very distinctive signatures on A_2, A_4 etc. Since our model expansion to A_4 only works to $k \sim 0.7\,h\,\mathrm{Mpc}^{-1}$, we explore this question numerically. In Fig. 11 bottom-right panel, we plot the supersample effect and amplitude effect such that they are degenerate at low k, while also adding the baryonic effect such that it is degenerate with change of amplitude up to $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$. We see from the Fig. 11 that above $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ the degeneracy is broken: the effects of σ_8 and δ_b are smaller compared to the effect of AGN feedback, which continuous to increase with k, because gas is being pushed out on small scales, suppressing all small-scale clustering. While this analysis is only restricted to a specific form of baryonic effects and is less robust than the other analyses in this paper, it suggests that one may be able to break the degeneracy between the baryonic effects, cosmological parameters such as amplitude, and supersample variance, using high k information.

6 DISCUSSION AND CONCLUSIONS

In this work, we propose a model of the matter power spectrum using the ZA power spectrum as the 2-halo term and even powers of k expansion of the 1-halo term, compensated on large scales to satisfy mass conservation, with coefficients calibrated on simulations. The leading order 1-halo term is k^0 term amplitude A_0, which in the halo model can be determined as a mass-dependent integral over the halo mass function. Simulations predict $A_0 \propto \sigma_8^{3.9}$, and the halo model is only able to reproduce this at low redshifts. The amplitude of A_0 is related to the cluster abundance method, where one counts clusters above a given mass, which also depends on the halo mass function and has a similarly steep dependence on σ_8. It is also related to Sunyaev-Zeldovich (SZ) power spectrum scaling, which is dominated by the 1-halo term and scales as σ_8^7 (Komatsu & Seljak 2002), because the SZ signal from individual clusters scales as $M^{5/3}$ rather than halo mass M and it is a projection over line of sight, leading to a steeper dependence on σ_8. Our analysis thus explicitly connects the cluster abundance method to the amplitude of the leading non-linear correction to the matter power spectrum, and shows the two use similar information. As a consequence, these two methods cannot be combined independently if the dominant errors are Poisson or large-scale structure fluctuations.

Using the first three coefficients of expansion, we accurately predict variations of basic cosmological parameters up to $k \sim 0.7\,h\,\mathrm{Mpc}^{-1}$, including amplitude σ_8, matter density Ω_m, Hubble parameter H_0, primordial slope n_s, equation of state w_0 and even neutrino mass $\sum m_\nu$. In all cases our model predicts well the BAO smoothing, a consequence of using the ZA rather than linear theory for the 2-halo term.

We present a very simple model for the covariance matrix of matter power spectrum (equation 34). We stress that the covariance matrix depends on the simulated volume both in linear and non-linear regimes, so a direct comparison between covariance matrices from different simulations needs to account for this. In this model the large-scale variance is dominated by the sampling variance, while on small scales where A_0 dominates the dominant term is the Poisson sampling of the haloes. Using the halo mass function of Tinker et al. (2008) to predict the latter gives about 20–30 per cent higher value than fitting with simulations of Li et al. (2014), which we consider good agreement given the inaccurate nature of halo mass function fits in the high mass regime. Using this value we show that our model gives remarkable agreement with the simulations of Blot et al. (2014), where 12 288 simulations of $656\,h^{-1}$ Mpc box size were run to construct a covariance matrix. We use our Poisson model to compute the convergence rate of the covariance matrix and find that simulated volumes of 500–5000 $(h^{-1}\,\mathrm{Gpc})^3$ are needed to converge at 1 per cent level. This explains why our model without any free parameters does not reproduce covariance matrix Harnois-Déraps & Pen (2012), because the total volume used in Harnois-Déraps & Pen (2012) was only 1.6 $(h^{-1}\,\mathrm{Gpc})^3$, and has thus not converged with high enough accuracy. Changing the parameter σ_{A_0}/A_0 from the predicted 0.09 to 0.15 we obtain perfect agreement.

Using this model we argue that most of the cosmological information about the amplitude is in A_0, which can determine the amplitude σ_8 to 0.2 per cent within 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume. The higher order coefficients A_2, A_4 etc. are less sensitive to σ_8 and have a larger variance. We discuss the supersample variance and argue that due to its origin as a curvature effect it differs from the amplitude rescaling and so it should be treated as a separate cosmological parameter with a prior given by the rms variance on the scale of the survey volume. If its degeneracy with the amplitude is not broken, then it approximately doubles the errors, so that σ_8 can be determined to 0.4 per cent within 1 $(h^{-1}\,\mathrm{Gpc})^3$ volume. Note that both of these errors are a lot smaller than the currently available constraints, which at best are at 4 per cent (Kilbinger et al. 2013): observational and modelling errors dominate the error budget at the moment, but future data sets may be able to reach the levels where supersampling variance or Poisson error will dominate (Yoo & Seljak 2012).

We also investigate the baryonic effects on the matter power spectrum. We argue that these should not change A_0 much because of the mass conservation. Indeed, comparison of our model to simulations of baryonic effects in van Daalen et al. (2011) suggests that A_0 is almost unchanged, while higher order coefficients change significantly, because baryonic effects redistribute gas and DM inside the haloes without changing the overall halo mass. We advocate that marginalizing over higher order expansion coefficients should immunize against baryonic

effects without much loss of information. We explore the degeneracy between the amplitude, supersample variance and baryonic effects, finding that it can be broken using information above $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$.

Our results suggest that analytic modelling of DM clustering provides important insights even in the era of large simulations. It offers a promising venue not only for an accurate power spectrum description, but also for the covariance matrix modelling, for optimal extraction of information from the data, and for description of baryonic effects. We have shown that in the context of covariance matrix calculations our model is likely to be more reliable than simulations with insufficient total volume. However, more work remains to be done before it can be applied to the weak lensing observations. For example, in this paper we focused on the DM clustering description in terms of its power spectrum. If one wants to apply the method to the weak lensing observations one needs to perform the line-of-sight projections of the model on to the weak lensing power spectrum $C_l^{\kappa\kappa}$, where κ is the convergence which can be written as a projection of the density along the line of sight. Projecting powers of k simply gives the same powers of l, so if the projection kernels are narrow, as would be the case for weak lensing tomography, the analysis remains essentially unchanged, except for the fact that weak lensing probes matter density rather than density perturbations, so convergence is also multiplied by an overall mean matter density. If the projection kernels are broad and there are significant contributions from nearby structures for which $k > 0.7\,h\,\mathrm{Mpc}^{-1}$ projects to a low l, then one needs to assess these effects and improve the model to account better for the high k contributions. Similarly, one also needs to project baryonic effects and covariance matrix. This programme is feasible and if implemented it will give a completely analytical description of the weak lensing power spectrum and its covariance matrix without any need to use simulations.

ACKNOWLEDGEMENTS

We thank Z. Vlah for extensive discussions and for providing the Zeldovich power spectrum code, F. Schmidt and M. Takada for useful comments, Y. Li for electronic form of the plots of Li et al. (2014), M. van Daalen for electronic form of the plots of van Daalen et al. (2011), L. Blot for providing us data from their simulations, and J. Harnois-Deraps and U. Pen for interpretation of their covariance matrix. US is supported in part by the NASA ATP grant NNX12AG71G. IM would like to thank the hospitality of LBNL.

REFERENCES

Baldauf T., Seljak U., Senatore L., Zaldarriaga M., 2011, J. Cosmol. Astropart. Phys., 10, 31
Baldauf T., Seljak U., Smith R. E., Hamaus N., Desjacques V., 2013, Phys. Rev. D, 88, 083507
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Bird S., Viel M., Haehnelt M. G., 2012, MNRAS, 420, 2551
Blot L., Corasaniti P. S., Alimi J.-M., Reverdy V., Rasera Y., 2014, preprint (arXiv:1406.2713)
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Fu L. et al., 2008, A&A, 479, 9
Hamilton A. J. S., Rimes C. D., Scoccimarro R., 2006, MNRAS, 371, 1188
Harnois-Déraps J., Pen U.-L., 2012, MNRAS, 423, 2288
Heitmann K., Higdon D., White M., Habib S., Williams B. J., Lawrence E., Wagner C., 2009, ApJ, 705, 156
Heitmann K., White M., Wagner C., Habib S., Higdon D., 2010, ApJ, 715, 104
Hoekstra H. et al., 2006, ApJ, 647, 116
Kilbinger M. et al., 2013, MNRAS, 430, 2200
Komatsu E., Seljak U., 2002, MNRAS, 336, 1256
Lawrence E., Heitmann K., White M., Higdon D., Wagner C., Habib S., Williams B., 2010, ApJ, 713, 1322
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Li Y., Hu W., Takada M., 2014, Phys. Rev. D, 89, 083519
Ma C., Fry J. N., 2000, ApJ, 543, 503
McClelland J., Silk J., 1977, ApJ, 217, 331
McCarthy I. G. et al., 2010, MNRAS, 406, 822
Massey R. et al., 2007, MNRAS, 376, 13
Matsubara T., 2008, Phys. Rev. D, 77, 063530
Mortonson M. J., Huterer D., Hu W., 2010, Phys. Rev. D, 82, 063004
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144
Percival W. J. et al., 2014, MNRAS, 439, 2531
Refregier A., 2003, ARA&A, 41, 645
Sato M., Hamana T., Takahashi R., Takada M., Yoshida N., Matsubara T., Sugiyama N., 2009, ApJ, 701, 945
Sato M., Takada M., Hamana T., Matsubara T., 2011, ApJ, 734, 76
Schneider P., Bartelmann M., 1995, MNRAS, 273, 475
Schrabback T. et al., 2010, A&A, 516, A63
Seljak U., 2000, MNRAS, 318, 203
Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020
Semboloni E., Hoekstra H., Schaye J., 2013, MNRAS, 434, 148
Sherwin B. D., Zaldarriaga M., 2012, Phys. Rev. D, 85, 103523
Slosar A., Hirata C., Seljak U., Ho S., Padmanabhan N., 2008, J. Cosmol. Astropart. Phys., 8, 31
Smith R. E. et al., 2003, MNRAS, 341, 1311

Takada M., Hu W., 2013, Phys. Rev. D, 87, 123504
Takada M., Spergel D. N., 2014, MNRAS, 441, 2456
Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
Taylor A. N., 1993, in Bouchet F., Lachieze-Rey M., eds, Cosmic Velocity Fields. Editions Frontieres, Gif-sur-Yvette, p. 585
Taylor A., Joachimi B., 2014, MNRAS, 442, 2728
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649
Yoo J., Seljak U., 2012, Phys. Rev. D, 86, 083504
Zel'dovich Y. B., 1970, A&A, 5, 84

This paper has been typeset from a TEX/LATEX file prepared by the author.

4

Paper3: Baryonic effects on weak-lensing two-point statistics

According to the Dark Energy Task Force (DETF; [4]), the two-point statistics of the weak gravitational lensing is amongst the most promising tools to do cosmology and derive strong constraints on the cosmological parameters. However, it is dominated by many systematics, both observational and theoretical. In this paper, we focussed on an extremely important source of systematic errors in the theory of weak lensing shear power spectrum - *baryonic effects*.

The theoretical modelling of the weak-lensing shear power spectrum is described in section 1.4.4 for a dark-matter only Universe, which is a fair approximation at large scales, which was the limit to the weak-lensing experiments in the past. The next generation surveys like Euclid, LSST etc., are pushing this limit far into the small scales such that baryonic contribution becomes important. In this paper, we built a model to incorporate baryonic contribution in the matter power spectrum, and by extension to the weak lensing shear power spectrum. Our model is based on the halo model where the baryonic contribution is sensitive to mainly two quantities: the halo mass function, and the radial density profiles of the halos.

Our baryonic model consists of four main ingredients: (i) hot intra cluster gas assumed to be in hydrostatic equilibrium; (ii) a stellar component dominated by a central galaxy whose mass is constrained by the abundance matching techniques; (iii) a feedback model that removes the gas from the halo as a function of its mass; and (iv) an adiabatically contracted dark-matter component. Incorporating these four components, we can reproduce hydrodynamical simulations for the radial profile of the halos, and the matter power spectrum.

We performed a cosmological parameter forecast for a Euclid like survey and

found that using weak-lensing alone, Euclid is expected to constrain the cosmological parameters to a very high accuracy. However, if the baryonic effects are not taken into account, it will bias the recovered values of the parameters and mislead the interpretations. On the other hand, if the baryonic effects are taken into account, all cosmological parameters can still be constrained with good accuracy along with the parameters of the baryonic model.

Role: I started this project under the supervision of Professor Romain Teyssier. I developed a code to compute the matter power spectrum in the halo model framework, assuming dark-matter only (DMO) Universe. I used Navarro Frenk White (NFW) radial density profiles for the dark-matter halos, and the mass-function from Tinker et al. 2008. The next task was to include the baryonic physics in the framework. I did this in three steps: (i) I included two baryonic components to the radial density profile; intra-cluster plasma, and a bright central galaxy (BCG), (ii) I changed the gas mass fraction from being a constant to be dependent on the mass of the halo regulated by a free parameter $M_{\rm crit}$, (iii) I included the adiabatic contraction of the dark-matter component due to the BCG. This comprises the full baryonic model. I compared the radial density profiles of this model to the simulations provided by Davide Martizzi, and found a remarkable agreement. Incorporating this prescription in the halo model framework, I computed the modified matter power spectrum with baryons, and studied its deviation from the DMO case for variable $M_{\rm crit}$ parameters. I wrote an extension of this code to compute the weak-lensing shear power spectrum for a given matter power spectrum. I studied the effects of baryons on weak lensing power spectrum for variable $M_{\rm crit}$ parameters. I studied the weak lensing observables in three redshift bins, and performed a tomographic analysis of the weak lensing shear power spectrum for a Euclid like survey. Using this code, I made a mock dataset with baryons, and included random noise in the observables. I ran two sets of 8 MCMC, each for a different $\ell_{\rm max}$. In the first set the model was dark-matter only, and in the second case, it included baryonic physics as well. To ensure convergence, each MCMC analysis was based on 16 individual chains.

This paper has been submitted to *Monthly Notices of the Royal Astronomical Society* (MNRAS), and is currently under review.
Arxiv: `http://arxiv.org/abs/1410.6826`

Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 28 October 2014 (MN LATEX style file v2.2)

Baryonic effects on weak-lensing two-point statistics and its cosmological implications

Irshad Mohammed*[1,2], Davide Martizzi[3], Romain Teyssier[2] and Adam Amara[4]
[1] *Physik-Institut, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
[2] *Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
[3] *Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA*
[4] *Institute for Astronomy, Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093, Zurich, Switzerland*

28 October 2014

ABSTRACT
We develop an extension of *the Halo Model* that describes analytically the corrections to the matter power spectrum due to the physics of baryons. We extend these corrections to the weak-lensing shear angular power spectrum. Within each halo, our baryonic model accounts for: 1) a central galaxy, the major stellar component whose properties are derived from abundance matching techniques; 2) a hot plasma in hydrostatic equilibrium and 3) an adiabatically-contracted dark matter component. This analytic approach allows us to compare our model to the dark-matter-only case. Our basic assumptions are tested against the hydrodynamical simulations of Martizzi et. al. (2014), with which a remarkable agreement is found. Our baryonic model has only one free parameter, $M_{\rm crit}$, the critical halo mass that marks the transition between feedback-dominated halos, mostly devoid of gas, and gas rich halos, in which AGN feedback effects become weaker. We explore the entire cosmological parameter space, using the angular power spectrum in three redshift bins as the observable, assuming a Euclid-like survey. We derive the corresponding constraints on the cosmological parameters, as well as the possible bias introduced by neglecting the effects of baryonic physics. We find that, up to ℓ_{max}=4000, baryonic physics plays very little role in the cosmological parameters estimation. However, if one goes up to ℓ_{max}=8000, the marginalized errors on the cosmological parameters can be significantly reduced, but neglecting baryonic physics can lead to bias in the recovered cosmological parameters up to 10σ. These biases are removed if one takes into account the main baryonic parameter, $M_{\rm crit}$, which can also be determined up to 1-2%, along with the other cosmological parameters.

Key words: Gravitational lensing: weak, methods: analytical, galaxies: halos, (cosmology:) cosmological parameters.

1 INTRODUCTION

The bending of light due to the presence of structures in its path is one very significant method to study the distribution of matter in the universe. The deflection is independent of the nature of the intervening matter, if it is dark or baryonic, and hence, this phenomenon, referred to as gravitational lensing, provides a unique tool to map the dark side of the universe. Under controlled systematics of the experiment, weak gravitational lensing, where the deflection of light rays are not significant enough to observe multiple images of the source but strong enough to deform the shape of the source, is a very powerful probe to study the nature of dark energy (Albrecht et al. 2006). The future sky surveys, like Euclid (Laureijs et al. 2011; Refregier 2009; Cimatti et al. 2009), are expected to provide maps of the sky with un-precedented accuracy and high resolution like never before (Amendola et al. 2013). It is an opportunity to employ the advantage of such high quality data to answer the most important questions in cosmology - the energy content of the universe, its dynamics, its evolution and the formation of structure. Weak gravitational lensing can be used as an ideal tool for such high quality data and can deliver, with sub-percent level accuracy, measurements of the main cosmological parameters.

The deformation of the shape of the observed galaxies due to the intervening matter is referred to as *shear*. This signal is very small, nearly 1% of the intrinsic ellipticity of

* irshad@physik.uzh.ch

the source galaxies, but can be measured statistically under the assumption that the intrinsic ellipticity of the background galaxies do not have a preferred direction. There are a number of interpretation of the two-point shear statistics based on dark matter only (collision-less) simulations which is a good approximation in the linear regime. However, at non-linear scales baryonic physics becomes important and can introduce a bias of 5 to 20 percent in the interpretation of the measurements, which in turn can introduce a bias in the cosmological constraints. So, in the era of precision cosmology, it is very important to quantify the effect of baryonic physics in the two-point shear statistics or the power spectrum.

Baryons account for nearly 20% of the matter content of the universe. Its distribution depends on the dark matter potential well, AGN feedback, supernovae, structure formation history and radiative cooling. Further baryonic distribution affects the matter power spectrum at small scales, which to the extension, affects the two point shear statistics. The effect of baryons on several statistics relevant for cosmology has been already studied by various authors. For instance, Stanek et al. (2009); Cui et al. (2012); Martizzi et al. (2014) and Cusworth et al. (2014) focused on the effects on the halo mass function. The effect of baryonic processes on the power spectrum and on the weak gravitational lensing shear signal has been studied too (White 2004; Zhan & Knox 2004; Jing et al. 2006; Rudd et al. 2008; Guillet et al. 2010; Semboloni et al. 2011; van Daalen et al. 2011; Reddick et al. 2014; Mohammed & Seljak 2014).

In most of the previous works (see references above), the approach was based on simulations, which suffer from finite volume and finite resolution effects, are performed using only one cosmology and baryonic model. They however capture the non-linear physics of gravitational collapse and the associated baryonic effects. In this work, we employ the halo model, an analytical approach, to build two-point shear statistics with and without baryons. This allows one to recover various different realizations of any cosmological models. We also compare our results with simulations at various stages to validate our main assumptions.

The outline of the paper is as follows: In section 2, we review the necessary concepts of the halo model and propose our baryonic model as a modification in the radial density profiles of the halos. We compare the model to simulations with AGN feedback models. We also review the modelling of shear power spectrum. We talk about the covariance matrix of the C_ℓ, Gaussian and non-Gaussian parts. In section 3, we make a comparison between the dark-matter-only model (DMO) and our baryonic model (BAR) and shows the behaviour of the baryonic correction as a function of our main AGN-feedback-parameter, $M_{\rm crit}$. We introduce our fiducial model and mock datasets to perform the likelihood analysis in section 4. In section 5, we talk about the cosmological implication of these baryonic corrections and the forecasts on the cosmological parameters, its accuracy and precision. Finally in section 6 we discuss the implications of our results and propose possible strategies for future works.

2 THEORETICAL MODEL - A SHORT REVIEW

We employ an analytic approach to model the effects of baryonic physics on the matter power spectrum and to the extension, on the shear power spectrum. The model has two broad parts: (*i*) the dark-matter-only model (DMO), and (*ii*) the modified model with baryonic physics (BAR). These two approaches modify the density profile of dark matter halos. We used the halo model (Seljak 2000; Peacock & Smith 2000; Ma & Fry 2000; Cooray & Sheth 2002) to construct the matter power spectrum based on the density profiles of halos of mass M and at redshift z.

2.1 The halo model

We employed the halo model (McClelland & Silk 1977; Seljak 2000; Ma & Fry 2000; Peacock & Smith 2000; Cooray & Sheth 2002) approach to calculate the matter power spectrum given the density profile of the halos. The halo model assumes all the matter in the universe to be in spherical halos with mass defined by a threshold density as:

$$M_{\triangle} = \frac{4}{3}\pi R^3_{\triangle} \triangle \bar{\rho}_m \tag{1}$$

where $M_{\triangle}$ is the mass of the halo and $R_{\triangle}$ is the boundary where the density of the halo drop to $\triangle$ times the mean matter density of the Universe, $\bar{\rho}_m$. We use $\triangle = 200$ throughout this paper, unless stated otherwise. We define the virial radius of the halo $R_{\rm vir}$ to be R_{200}.

In this framework, the matter power spectrum can be split into two parts:

$$P(k) = P_{1h}(k) + P_{2h}(k), \tag{2}$$

where, the two terms on the right hand side correspond to 1-halo term, describing the correlation between dark matter particles within the halo and 2-halo term which describes the halo-halo correlation respectively. These terms are given by

$$P_{1h} = \int d\nu (f_{dm} + f_{gas}(\nu)) f(\nu) \frac{M}{\rho} |u(k|\nu)|^2, \tag{3}$$

$$P_{2h} = \left(f_0 b_0 + \int d\nu (f_{dm} + f_{gas}(\nu)) f(\nu) u(k|\nu) b(\nu) \right)^2 P_{\rm lin}(k), \tag{4}$$

where, M is the mass of the halo and $\nu = \delta_c/\sigma(M,z)$ with $\delta_c = 1.686$. The term $f(\nu)$ is the functional form of the mass function and we used the fitting formula from Tinker et al. (2008). The term $b(\nu)$ resembles the bias in the dark matter halos and we used the fitting formula in (Tinker et al. 2010). To fulfill the underlying assumptions of the halo model, these two functional forms, f_ν and b_ν have to be expressed as in the following relations:

$$\int_0^\infty f(\nu) d\nu = 1 \tag{5}$$

$$\int_0^\infty f(\nu) b(\nu) d\nu = 1 \tag{6}$$

However, assuming a lower mass cut corresponding to $\nu_{\rm min}$, we introduce new background factors f_0 and b_0 such that:

$$f_0 + \int_{\nu_{\rm min}}^{\infty} (f_{dm} + f_{gas}(\nu)) f(\nu) d\nu = 1 \tag{7}$$

$$f_0 b_0 + \int_{\nu_{\rm min}}^{\infty} (f_{dm} + f_{gas}(\nu)) f(\nu) b(\nu) d\nu = 1 \tag{8}$$

Additionally, the term $f_{dm} + f_{gas} = 1$ for simpler models like no feedback, but for more exotic models, like with AGN feedback or including other baryonic physics, this term may deviate from unity. This will be more useful as explained in section 2.3

We used the Eisenstein & Hu (1998, 1999) transfer function calculations to account for the linear matter power spectrum term, $P_{lin}(k)$. The term $u(k|M)$ is the Fourier transform of the normalized density profile and is given by,

$$u(k|M) = \frac{4\pi}{M} \int_0^{R_{\rm vir}} dr\ r^2\ \rho(r|M)\ \frac{\sin(kr)}{kr}. \tag{9}$$

where, $\rho(r|M)$ is the density profile of the halo of mass M. The function $u(k|M)$ is normalised such that $u(k = 0|M) = 1$ .The dispersion of the smoothed density field, $\sigma(M, z)$, is given by,

$$\sigma^2(M, z) = \frac{1}{2\pi^2} \int P_{\rm lin}(k) k^2 |\tilde{W}(R, k)|^2 dk, \tag{10}$$

where, $\tilde{W}(R, k)$ is the Fourier transform of top-hat filtering function and given by,

$$\tilde{W}(R, k) = 3 \frac{\sin(kR) - kR\cos(kR)}{(kR)^3} \tag{11}$$

This framework of the halo model is applied to both DMO and BAR model which, differ in the halo density profiles and normalization of the mass function. The following two sections explains the corresponding profiles.

2.2 Dark matter only

We started with the radial density profile of dark matter halos given by the functional form:

$$\rho(r|M) = \frac{\rho_s}{(r/R_s)^{\alpha}(1 + r/R_s)^{\beta}}, \tag{12}$$

where, R_s is the characteristic radius given by the concentration parameter (c) and the virial radius of the halo (r_{vir}) as $c = R_{\rm vir}/R_s$. We used the two parameters α and β to be 1 and 2 respectively, corresponding to the Navarro-Frenk-White (NFW) profile (Navarro et al. 1997). The characteristic density ρ_s which is strongly degenerate with R_s and also proportional to the critical density of the Universe when the halo was formed. So, the NFW profile for dark matter halos is completely described by its concentration.

The concentration parameter c gives the information about the environment or the mean background density during the formation of the halo. A number of N-body simulations (Navarro et al. 1997; Avila-Reese et al. 1999; Jing 2000; Bullock et al. 2001; Eke et al. 2001; Zhao et al. 2003; Neto et al. 2007; Macciò et al. 2007; Duffy et al. 2008; Gao et al. 2008; Dutton & Macciò 2014) has prescribed various power laws between mass of the halo (M) and its concentration parameter c at redshift z. We used the fitting formula given in (Muñoz-Cuartas et al. 2011):

$$\log(c) = a(z) \log(M_{\rm vir}/[h^{-1} M_{\odot}]) + b(z) \tag{13}$$

where,

$$a(z) = \omega z - m \tag{14}$$

and

$$b(z) = \frac{\alpha}{z + \gamma} + \frac{\beta}{(z + \gamma)^2} \tag{15}$$

The fitting parameters $\omega,\ m,\ \alpha,\ \beta$ and γ are 0.029, 0.097, -110.001, 2469.720 and 16.885 respectively. Figure 1 (top-left panel) shows the behaviour of the concentration parameter as function of halo mass at different redshifts. There is an anti-correlation between the mass of the halo and its concentration. Also for a given halo mass, the concentration decreases with redshift. We limit the minimum concentration to 4 (dashed line in figure 1 upper-left panel). This is because the higher mass halos did not reach there maximum formation efficiency redshift and will reach it in future. So, on an average, there concentration must not be less than a few. A very recent study from Dutton & Macciò (2014) shows that this behaviour is consistent and the minimum concentration is very close to 4.

2.3 A baryonic model

Our baryonic model accounts within each halo for: 1) a central galaxy, the major stellar component whose properties are derived from abundance matching techniques; 2) a hot plasma in hydrostatic equilibrium and 3) an adiabatically-contracted (AC) dark matter component. This analytic approach allows us to compare our model to the DMO case. Apart from the normalization of the mass function, there is only one term that is affected by these baryonic components and is the density profile of the halo, which no longer follows the NFW profile. We can write the modified NFW (BAR) profile as:

$$\rho_{\rm BAR}(r|M) = f_{\rm dm} \rho^{\rm AC}_{\rm NFW}(r) + \rho_{\rm BCG}(r) + f_{\rm gas}(M) \rho_{\rm gas}(r), \tag{16}$$

we discuss each of these terms in more details.

2.3.1 Stellar component

We used the fitting function from Moster et al. (2013) based on abundance matching to map the stellar mass of the central galaxy $M_{\rm CentralGalaxy}$ (BCG), which is the major component of stellar mass in a cluster, to the mass of the halo ($M_{\rm halo}$). Figure 1 (top-right panel) shows the mapping between halo mass and stellar mass fraction associated to the central galaxy for a variety of redshifts. The relation has a

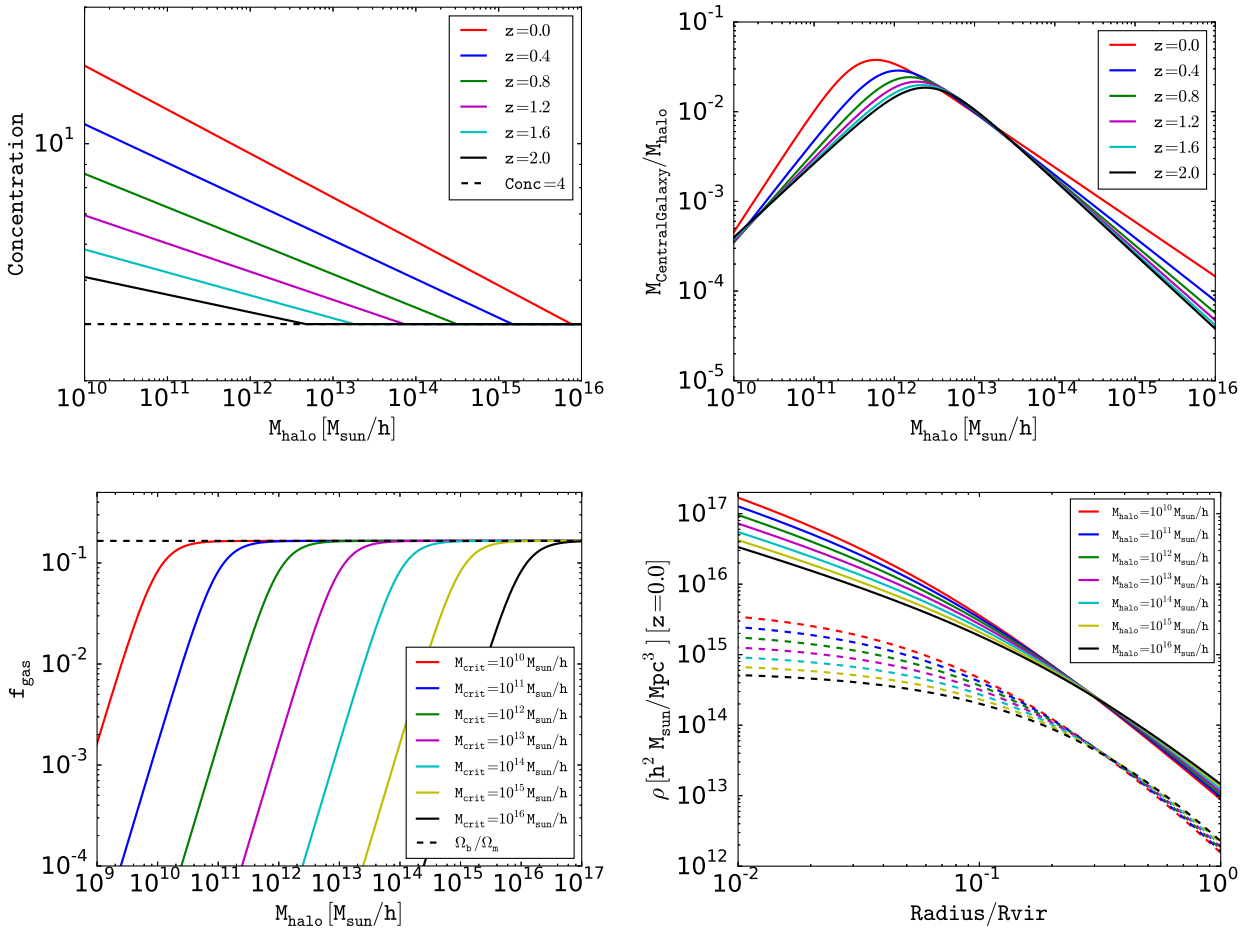

Figure 1. Top left: Concentration parameter as a function of halo mass for variable redshift. Top right: Mass of the central galaxy as a function of halo mass for variable redshift. Bottom left: Gas mass fraction as a function of halo mass for variable $M_{\rm crit}$. Bottom right: Density profile for NFW (solid lines) and intra-cluster gas (dashed lines) for different halo masses at redshift 0.

positive slope for low mass halos, however, at about the size of the Milky way halo, the slope turns negative. At this peak, the central galaxy stellar mass contributes about 4-5 % of the total mass of the halo. Also this peak shifts to higher masses for higher redshifts but contributes lower fraction.

The actual distribution of stellar mass in galaxy groups and clusters can be quite complex. The total stellar mass budget can be decomposed in 3 components: satellite galaxies, Brightest Cluster Galaxy (BCG, the massive elliptical galaxy dominating the cluster centre) and Intra-cluster Light (ICL, an extended stellar halo surrounding the BCG). The BCG and ICL represent $\sim$ 40 % of the mass in clusters, with this ratio decreasing with total cluster mass (Gonzalez et al. 2007). However, BCG+ICL dominate the inner part of the cluster and constitute $\sim$ 70% of the total stellar mass within 0.1 R_{200}. This fact is particularly relevant for computing the effect of baryon condensation on the dark matter profiles (see Subsection 2.3.3). The BCG+ICL component is usually modelled using superimposition of fitting functions, typically multiple Sersic profiles. Given that we are not interested in detailed modelling of the stellar distribution, we consider a simplified model for the BCG+ICL.

we adopted a radial density profile for BCG, where the enclosed mass goes linearly with the radius,

$$M_\star(< r) = M_{\rm CentralGalaxy}\frac{r}{2R_{1/2}} \tag{17}$$

this gives,

$$\rho(r) = \frac{M_{\rm CentralGalaxy}}{8\pi R_{1/2} r^2},\ r < 2R_{1/2} \tag{18}$$

where, $R_{1/2}$ is the half mass radius. We use $R_{1/2} = 0.015R_{\rm vir}$ which is a good fit to the observations (Kravtsov et al. 2014). We forced the density profile to drop exponentially after $2R_{1/2}$.

2.3.2 *Intra-cluster plasma*

The major component of the baryonic matter in a galaxy cluster is the hot intra-cluster gas. It is mainly ionized hydrogen at very high temperature and low density. This plasma radiates in X-rays and can safely be assumed to be in hydrostatic equilibrium. We assume this gas distribution in the halo according the hydrostatic equilibrium equations given in Martizzi et al. (2013),

$$\rho(x) = \rho_0 \left[\frac{\ln(1+x)}{x}\right]^{\frac{1}{\Gamma-1}} \tag{19}$$

where, x is the distance from the centre of the halo in unit of scale radius R_s. The effective polytropic index Γ is given by,

$$\Gamma = 1 + \frac{(1+x_{eq})\ln(1+x_{eq}) - x_{eq}}{(1+3x_{eq})\ln(1+x_{eq})} \tag{20}$$

where, $x_{eq} = c/\sqrt{5}$. Figure 1 (bottom-right in dashed lines) shows the density profile of the hot gas for variable halo masses at redshift 0 and also shows the comparison to the NFW profile (solid lines). For $x > x_{eq}$, the gas density profiles follows the NFW profile, however, it approaches a nearly constant values near the centre of the halo.

The normalization of the gas density profile, ρ_0, is fixed by the gas fraction $f_{\rm gas}$. if we assume no feedback from the baryonic component of the halo, this number can be a constant, however, many hydrodynamical simulations (Read & Gilmore 2005; Dehnen 2005; Mashchenko et al. 2006; Pontzen & Governato 2012; Teyssier et al. 2013; Martizzi et al. 2013) shows signatures of the expulsion of gas from the halo. This expulsion is stronger in low mass halos than the high mass halos. So the low mass halos are generally deficit in this hot plasma component. Following the same physical motivation, we used the gas mass fraction of the halo to be the function of the mass of the halo following the parametric form:

$$f_{\rm gas}(M_{\rm halo}) = \frac{\Omega_b/\Omega_m}{1 + \left(\frac{M_{\rm crit}}{M_{\rm halo}}\right)^{\beta}} \tag{21}$$

where, $M_{\rm crit}$ is a free parameter and β is fixed to 2. This parameter controls the gas fraction in halos of different mass. A higher value for $M_{\rm crit}$ represents less gas in the halo up to higher halo masses. This parameter can also be interpreted as the control sequence for AGN feedback. Figure 1 (bottom-left panel) shows the variation of $f_{\rm gas}$ with halo mass for variety of $M_{\rm crit}$. We chose $M_{\rm crit} = 10^{13}h^{-1}M_{\odot}$ as the most realistic model. In this case, all halos with mass lower than $\sim 2\times10^{12}h^{-1}M_{\odot}$ have expelled all their gas to the background (outside the $R_{\rm vir}$) and all halos with mass larger than $\sim 2\times10^{13}h^{-1}M_{\odot}$ have all their gas inside the halo. The intermediate mass halos have a very smooth transition from no gas to all gas inside the halo. This behaviour matches well with recent study from Schaller et al. (2014). We studied this case in detail for all its cosmological implications at different scales. We also studied one optimistic[1] model, where the feedback is not as strong as in our realistic model, with $M_{\rm crit} = 10^{12}h^{-1}M_{\odot}$.

[1] Optimistic in the sense of less AGN feedback that makes the baryonic corrections less troublesome

Type	H_0	σ_8	$n_{\rm s}$	Ω_Λ	$\Omega_{\rm m}$	$\Omega_{\rm b}$
DMO	70.4	0.809	0.963	0.728	0.272	-
BAR	70.4	0.809	0.963	0.728	0.272	0.045

Table 1. Cosmological parameters adopted in our simulations.

Type	$m_{\rm cdm}$ [10^8 $M_\odot$/h]	$m_{\rm gas}$ [10^7 $M_\odot$/h]	$\Delta x_{\rm min}$ [kpc/h]
Original box	15.5	n.a.	2.14
DMO zoom-in	1.94	n.a.	1.07
BAR zoom-in	1.62	3.22	1.07

Table 2. Mass resolution for dark matter particles, gas cells and star particles, and spatial resolution (in physical units) for our simulations.

2.3.3 Adiabatic contraction

In the DMO model, we adopted the NFW profile for the distribution of dark matter in the halo which is nearly scale-free and completely described by the concentration parameter. However, in the presence of baryons, the dark matter component follows NFW only in the outskirts of the halo, but in the very centre the dark matter profile becomes steeper and deviates from pure a NFW profile. This is because the baryons, which are dominant in the centre of the halo, drag some extra matter from the surrounding towards the centre making the dark matter profile steeper towards the centre. The total distribution of matter is expected to dynamically respond to the condensation of baryons at the centre of the halo in a way that approximately conserves the value the adiabatic "invariant" $R \times M(R)$, where R is the distance from the halo centre and $M(R)$ is the mass enclosed in a sphere of radius R (Blumenthal et al. 1986; Gnedin et al. 2004). We adopted a simplified model for this effect following the appendix of Teyssier et al. (2011) where this adiabatic contraction (AC) of the dark matter profile is solely governed by the central galactic disk.

2.4 Comparison with simulations

We consider data from a set of cosmological re-simulations performed with the RAMSES code Teyssier (2002). These simulations are part of a larger set recently used by Martizzi et al. (2014) to study the baryonic effects on the halo mass function. Thanks to the adaptive mesh refinement capability of the RAMSES code, the resolution achieved in these simulations is sufficient to study the properties of low redshift BCGs.

In these calculations, the cosmological parameters are: matter density parameter $\Omega_{\rm m} = 0.272$, cosmological constant density parameter $\Omega_\Lambda = 0.728$, baryonic matter density parameter $\Omega_{\rm b} = 0.045$, power spectrum normalization $\sigma_8 = 0.809$, primordial power spectrum index $n_{\rm s} = 0.963$ and Hubble constant $H_0 = 70.4$ km/s/Mpc (Table 1). We generated initial conditions for the simulations using the Eisenstein & Hu (1998) transfer function and

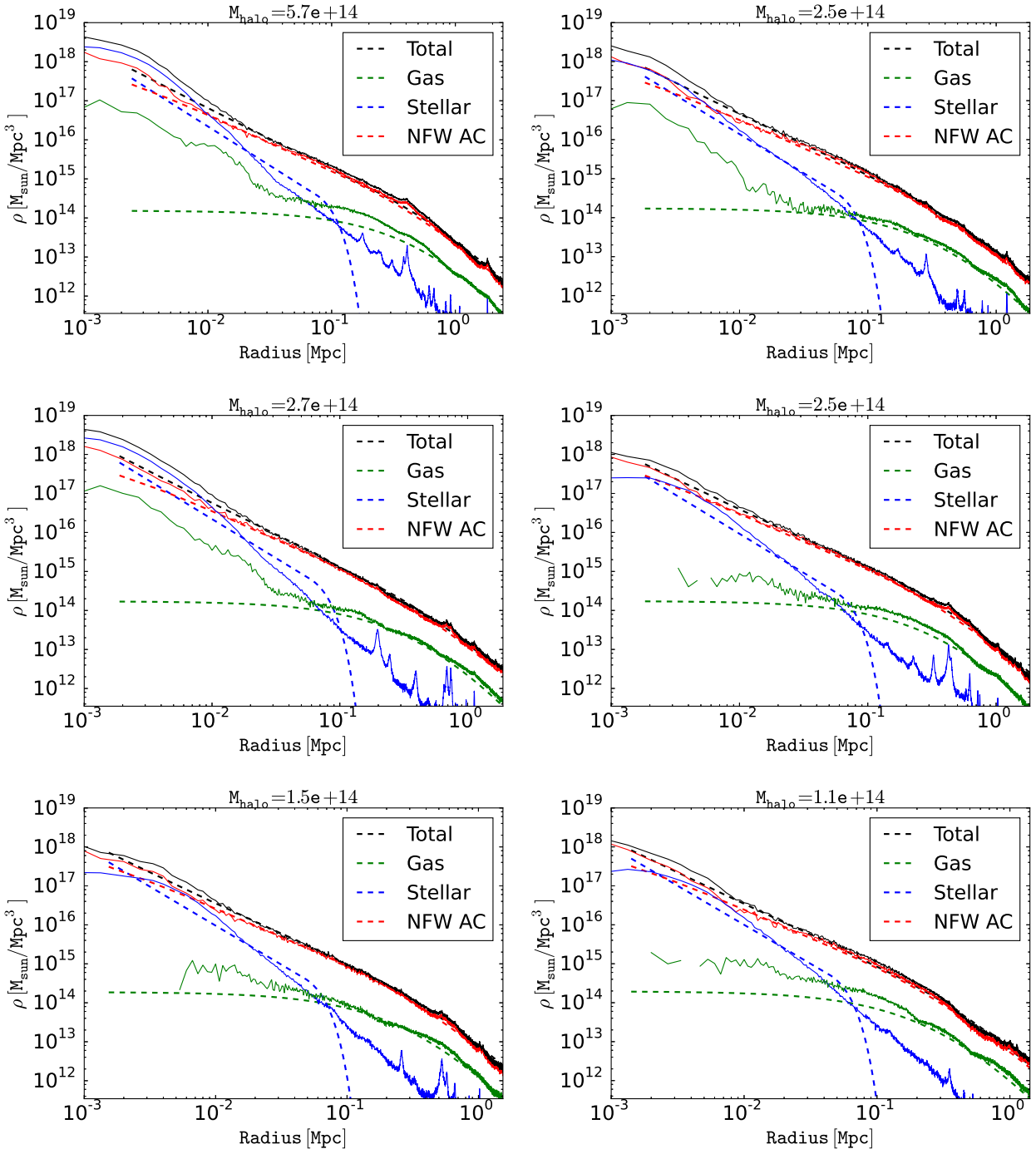

Figure 2. A comparison of our model density profiles (dashed lines) with hydrodynamical simulations of Martizzi et. al. 2014 (solid lines). There is a remarkable agreement, except at the very centre of the halo.

the GRAFIC++ code[2], based on the original GRAFIC code (Bertschinger 2001). These simulations come in two flavours: DMO (dark matter only) which only follow the evolution of dark matter, BAR which include baryons and galaxy formation prescriptions.

[2] http://sourceforge.net/projects/grafic/

The technique we adopted to perform the zoom-ins is described in the following. First, we ran a dark matter only simulation with particle mass $m_{\rm cdm} = 1.55 \times 10^9$ M$_\odot$/h and box size 144 Mpc/h. The initial level of refinement was $\ell = 9$ (512^3), but as the simulation evolved more levels of refinement were allowed. At redshift $z = 0$ the grid was refined down to a maximum level $\ell_{\rm max} = 16$. Subsequently, we ran apply the AdaptaHOP algorithm Aubert et al. (2004) to

identify the position and masses of dark matter halos. We selected 51 halos whose *total* masses lie $M_{\rm tot} > 10^{14}$ $M_\odot$ and whose neighbouring halos do not have masses larger than $M/2$ within a spherical region of five times their virial radius. We determined that only 25 of these clusters are relaxed. High resolution initial conditions were extracted for each of the 51 halos and were used to run zoom-in re-simulations. Three different re-simulations per halo have been performed: (I) including dark matter and neglecting baryons, (II) including dark matter, baryons and stellar feedback, (III) including baryons, stellar feedback and AGN feedback. In this paper we focus on cases (I) and (III), labelled DMO and BAR, respectively.

In the DMO re-simulation, the dark matter particle mass is $m_{\rm cdm} = 1.94\times10^8$ $M_\odot$/h. In the BAR re-simulations, the dark matter particle mass is $m_{\rm cdm} = 1.62 \times 10^8$ $M_\odot$/h, while the baryon resolution element has a mass of $m_{\rm gas} = 3.22 \times 10^7$ $M_\odot$. The maximum refinement level was set to $\ell = 17$, corresponding to a minimum cell size $\Delta x_{\rm min} = L/2^{\ell_{\rm max}} \simeq 1.07$ kpc/h. The grid was dynamically refined using a quasi-Lagrangian approach: when the dark matter or baryonic mass in a cell reaches 8 times the initial mass resolution, it is split into 8 children cells. Table 2 summarizes the particle mass and spatial resolution achieved in the simulations.

The physical prescription implemented in the code to perform the BAR simulations is here briefly described. In RAMSES gas dynamics is solved via a second-order unsplit Godunov scheme (Teyssier 2002) based on different Riemann solvers (we adopted the HLLC solver) and the MinMod slope limiter. The gas is described by perfect gas equation of state (EOS) with polytropic index $\gamma = 5/3$. Gas cooling is modelled with the Sutherland & Dopita (1993) cooling function which accounts for H, He and metals. Star formation and supernovae feedback ("delayed cooling" scheme, Stinson et al. (2006)) and metal enrichment have been included in the calculations. AGN feedback has been included too, using a method inspired by the Booth & Schaye (2009) model. In this scheme, super-massive black holes (SMBHs) are modeled as sink particles and AGN feedback is provided in form of thermal energy injected in a sphere surrounding each SMBH. More details about the AGN feedback scheme and about the tuning of the galaxy formation prescriptions can be found in Teyssier et al. (2011) and Martizzi et al. (2012).

Figure 2 shows the comparison between the dark matter, gas, stellar and total mass density profiles of 6 halos in the Martizzi et al. (2014) catalogue and the mass model described in Section 2.1. The model for the adiabatically contracted dark matter profile (red dashed lines) fits well the simulations down to scales ~ 10 kpc. The model for the Intra-cluster plasma (green dashed lines) fits well the results of the simulations down to scales ~ 50 kpc. The relation between mass of the central galaxy and that of the halo has a lot of scatter. So, to compare with simulations we use the stellar mass from the simulation itself for the given halo, which define the normalisation of our stellar model. The model (blue dashed lines) is a good fit to the results of the simulations except in the outskirts. This is expected since the data from the simulations include BCG, ICL and satellite galaxies. However, the model is constructed in such a way that the stellar mass expected from abundance matching is associated to the central regions of the halos. The overall result is that the model for the total mass (black dashed lines) provides an excellent match to the results of cosmological simulations down to a scale of ~ 10 kpc. Therefore we conclude that the mass model is good enough to be adopted for the purposes of this paper.

2.5 From $P(k)$ to $C(\ell)$

In this section we develop the mapping from 3D matter power spectrum $P(k, z)$ to the 2D projected shear angular power spectrum C_ℓ following the theoretical framework explained in Takada & Jain (2009).

The distortion of the source shape due to weak gravitational lensing can be quantified with two quantities: shear γ and convergence κ. The convergence κ is the local isotropic part of the deformation matrix and can be expressed as:

$$\kappa(\vec{\theta}) = \frac{1}{2}\vec{\nabla}.\vec{\alpha}(\vec{\theta}) \tag{22}$$

where, α is the deflection angle. If we know the redshift of the source galaxies, additional information can be gained by dividing the sources in different redshift bins. This process is referred to as lensing tomography and is very useful to gain extra constraints on cosmology from the evolution of the weak lensing power spectrum (Hu 1999; Huterer 2002; Takada & Jain 2004). In cosmological context, the convergence field can be expressed as the weighted projection of the mass distribution integrated along the line of sight in the ith redshift bin,

$$\kappa_i(\vec{\theta}) = \int_0^{\chi_H} g_i(\chi)\delta(\chi\vec{\theta}, \chi)d\chi, \tag{23}$$

where, δ is the total 3 dimensional matter overdensity, χ is the comoving distance and χ_H is the comoving distance to the horizon. For a complete review see Mellier (1999); Bartelmann & Schneider (2001); Schneider (2006). The lensing weights $g_i(\chi)$ in the ith redshift bin with comoving distance range between χ_i and χ_{i+1} are given by:

$$g_i(\chi) = \begin{cases} \frac{g_0}{\bar{n}_i}\frac{\chi}{a(\chi)}\int_{\chi_i}^{\chi_{i+1}} n_s(\chi')\frac{dz}{d\chi'}\frac{(\chi'-\chi)}{\chi'}d\chi', & \chi \le \chi_{i+1} \\ 0, & \chi > \chi_{i+1} \end{cases} \tag{24}$$

where, $a(\chi)$ is the scale factor at comoving distance χ. Also,

$$g_0 = \frac{3}{2}\frac{\Omega_m}{H_0^2} \tag{25}$$

and,

$$\bar{n}_i = \int_{\chi_i}^{\chi_{i+1}} n_s(\chi(z))\frac{dz}{d\chi'}d\chi'. \tag{26}$$

where, $n_s(\chi(z))$ is the distribution of sources in redshift. We assume a source distribution along the line of sight of the form:

$$n_s(z) = n_0 \times 4z^2 \exp\left(-\frac{z}{z_0}\right) \tag{27}$$

with $n_0 = 1.18 \times 10^9$ per unit steradian and z_0 is fixed such that the corresponding projected source density n_g resembles the experiment, like Euclid etc.

$$\int_0^\infty n_s(z)dz = \bar{n}_g. \tag{28}$$

For Euclid like survey, we choose z_0 such that $\bar{n}_g = 50$ sources per arcmin^{-2} (Hoekstra & Jain 2008).

Finally the shear power spectrum between redshift bins i and j can be computed as:

$$C_{ij}(\ell) = \int_0^{\chi_H} \frac{g_i(\chi)g_j(\chi)}{\chi^2} P\left(\frac{\ell}{\chi}, \chi\right) d\chi \tag{29}$$

where, P is the 3D matter power spectrum calculated using the halo model framework as described in section 2.1. Larger ℓ corresponds to the smaller scale and the large contribution of C_ℓ at higher ℓ comes from non-linear clustering.

We divided the big cosmological volume into 3 redshift bins with boundaries: 0.01, 0.8, 1.5 and 4.0; so we calculated total 6 convergence cross-spectra (3 auto-spectra and 3 cross-spectra).

The auto-spectra is contaminated by the intrinsic ellipticity noise and assuming its distribution to be completely uncorrelated to different source galaxies, the observed power spectrum $C_{ij}^{\rm obs}(\ell)$ is given by,

$$C_{ij}^{\rm obs}(\ell) = C_{ij}(\ell) + \delta_{ij}\frac{\sigma_\epsilon^2}{\bar{n}_i}, \tag{30}$$

we choose $\sigma_\epsilon = 0.33$ which is the RMS intrinsic ellipticity. The cross spectra is not contaminated by shot noise.

The covariance matrix of C_ℓ has two contributions: Gaussian and non-Gaussian (NG). In this work we only consider the Gaussian contribution to the covariance matrix which is given by the following expression,

$$\mathrm{Cov}_{ij,mn}(\ell,\ell') = \frac{\delta_{\ell\ell'}}{\Delta\ell(2\ell+1)\mathrm{f_{sky}}} \times \left(C_{im}^{\rm obs}(\ell)C_{jn}^{\rm obs}(\ell) + C_{in}^{\rm obs}(\ell)C_{jm}^{\rm obs}(\ell)\right), \tag{31}$$

where, $\Delta\ell$ is the bin width of the ℓ and f_{sky} is the sky fraction for the targeted experiment. This term is dominated by cosmic variance for lower ℓ and shot noise for higher ℓ, however, for large number of sources, as in case of Euclid, and larger size of bins ($\Delta\ell$) towards higher end of ℓ, the shot noise can be significantly reduced.

The NG contribution to the covariance matrix of C_ℓ is rather complicated to calculate. It gives the correlation between different ℓ. At the matter power spectrum level, this term depends on the matter trispectrum. To compute the NG covariance to lensing, we need to integrate the trispectrum in redshift and angle on the sky and then compute this quantity for various ℓ and ℓ'. So this is a 4D calculation of trispectrum which is computationally very expensive. Yoo & Seljak (2012) shows that these NG correction to the covariance becomes significant for ℓ of few thousand and Cooray & Sheth (2002) shows that neglecting this will introduce the bias in the cosmological parameters up to 20 %. In this work, we are not taking into account these corrections and we are doing our analysis for different ℓ_{max}: 1000, 2000, 3000, 4000, 5000, 6000, 8000, 10000 and 20000. We will discuss more about the NG covariance in section 6.3.

3 COMPARING BAR AND DMO MODEL

In this section we try to draw a comparison between the baryonic model (BAR) and the dark-matter only (DMO) model. We would like to establish an understanding of the scales where the baryonic corrections become important and how these scales changes with redshift and the only free parameter, $M_{\rm crit}$.

Figure 3 (top-left panel) shows the relative differences between the BAR and DMO predictions for the matter power spectrum, also referred as *boost* in this article. There is only one free parameter of the baryonic model, $M_{\rm crit}$ which regulates the amount of AGN feedback and which is introduced in section 2.3.2. The overall shape of the deviation is similar in all cases for various $M_{\rm crit}$ and redshifts: the BAR model follows the DMO model for large scales, suffers a deficit in power at intermediate scales due to flatter gas profile compared to the NFW profile and finally the power shoots up due to the central stellar component. Also without adiabatic contraction (AC) the raise in the matter power spectrum occurs at very small scales, but including AC effect this raise can be seen at comparatively lower k or larger scales. This is because AC makes the profile steeper in the centre and shallower in the outskirts.

At redshift 0 (top-left panel of figure 3), the baryonic correction starts showing up (more than 1%) at $k \sim 5\ h/Mpc$ for models with negligible AGN feedback (lower $M_{\rm crit}$), whereas for more extreme AGN feedback models (higher $M_{\rm crit}$) this correction is important at much larger scales like $k \sim 0.1\ h/Mpc$. In our fiducial BAR model with $M_{\rm crit} = 10^{13}h^{-1}M_\odot$, the baryonic effects become significant, i.e., more than 1 percent, at $k \sim 0.5\ h/Mpc$. The maximum dip in the intermediate scales vary for different $M_{\rm crit}$; for the most extreme models where AGN feedback can push all the gas out of the halo, this dip is nearly the cosmic baryon fraction, Ω_b/Ω_m. However, for more a realistic model ($M_{\rm crit} = 10^{13}h^{-1}M_\odot$) this dip is nearly 7-8%. For more optimistic models like $M_{\rm crit} = 10^{12}h^{-1}M_\odot$, this dip is even smaller, nearly 4-5%. Therefore, we can conclude the more extreme AGN feedback models triggers the deviation of matter power spectrum from DMO model at larger scales and also the dip in the power at intermediate scales can be as large as the cosmic baryon fraction in case where all the gas are pulled out by the AGN feedback, however, for more realistic and optimistic models, the deviation starts at relatively small scales and also the maximum dip is comparatively smaller.

Figure 3 (top-right panel) shows the same quantity for a fixed $M_{\rm crit} = 10^{13}h^{-1}M_\odot$ at different redshifts. If we go to higher redshift, the overall shape of the deviation of the BAR matter power spectrum from the prediction of the DMO model (boost) is nearly the same as at redshift zero, however, the scales and the maximum dip amplitude at various redshifts change. We see that at higher redshifts, the dip

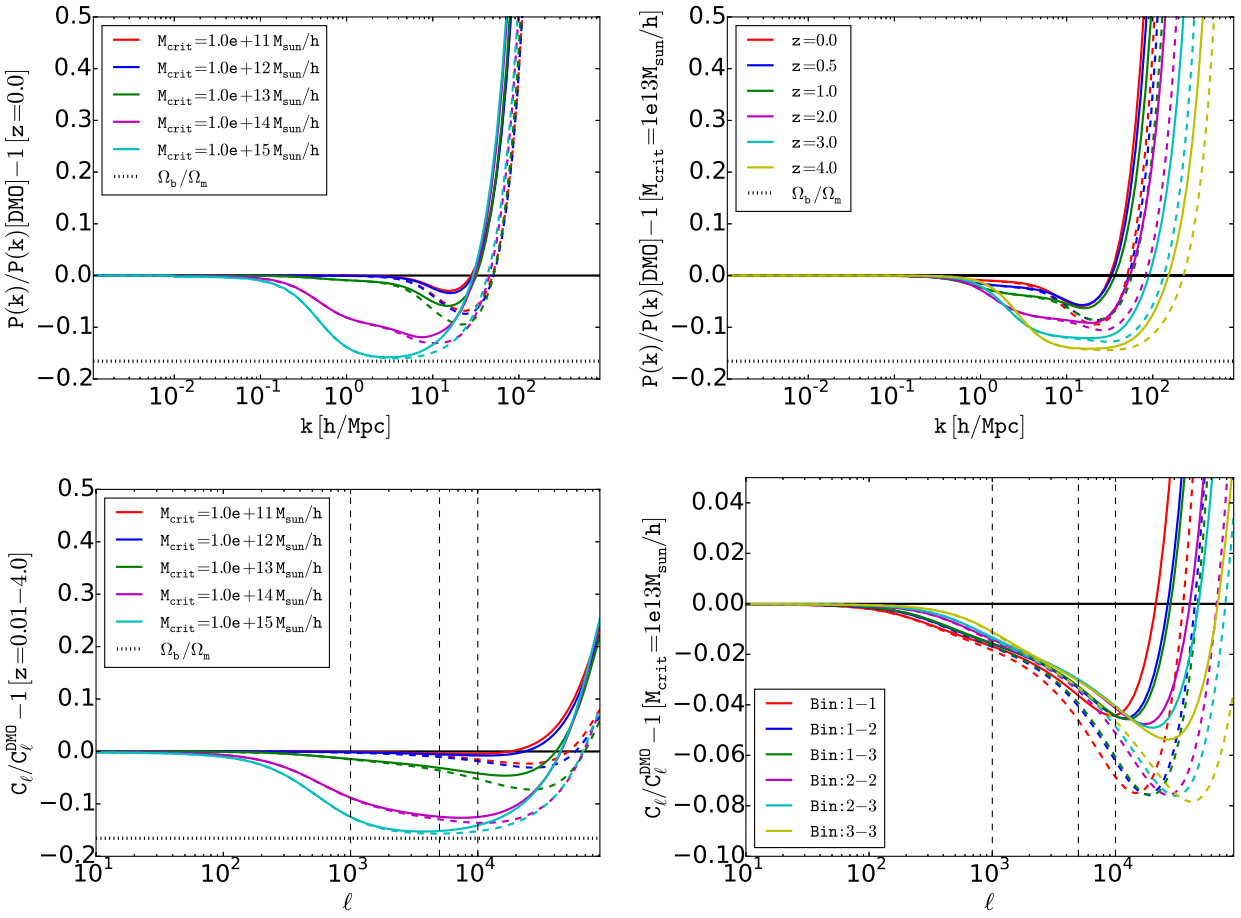

Figure 3. Top row: Relative deviation of the matter power spectrum predicted by the BAR model from the DMO model predictions as a function of k for different $M_{\rm crit}$ at redshift zero (left) and for fixed $M_{\rm crit} = 10^{13}h^{-1}M_{\odot}$ and different redshifts (right). Bottom row: Relative deviation of the shear power spectrum (C_ℓ) predicted by the BAR model from DMO model predictions for different $M_{\rm crit}$ in one big redshift bin (left) and for three tomographic redshift bins and fixed $M_{\rm crit} = 10^{13}h^{-1}M_{\odot}$ (right). Dashed lines are the calculations without adiabatic contraction (AC) and solid lines with adiabatic contraction (AC). The horizontal dashed line shows the cosmic baryon fraction.

starts to trigger at larger scales and also the maximum dip converge to the cosmic baryon fraction.

In figure 3 (bottom-left panel), the baryonic correction to C_ℓ is shown in one big redshift bin ($z = 0.01 - 4.0$). Here, the shear power spectrum starts to deviate from DMO predictions at about $\ell = 100$ for the most extreme AGN feedback models and at ℓ of about several thousands for models with weak AGN feedback. For our realistic model (green curve), this deviation occurs at about $\ell \sim 700$. The maximum dip in power is very similar to that of the matter power spectrum explained above. It is worth noticing that for $\ell = 10000$ the deviation is very significant for the realistic model ($M_{\rm crit} = 10^{13}h^{-1}M_{\odot}$), however, it is negligible for the optimistic model ($M_{\rm crit} = 10^{12}h^{-1}M_{\odot}$). Because these are the cases that we study in our likelihood analysis, we will show in section 5 that this behaviour is consistent with the cosmological parameter estimation with these models.

4 FIDUCIAL MODEL AND MOCK DATASETS

In this section, we would like to mention two factors that are quite important for our experiments - fiducial parameters and mock datasets. The fiducial parameters assumed in this work, particularly about cosmology, baryonic model and Euclid mission, are very standard. Also the mock datasets generated are correctly contaminated with random noise. Following are the key numbers and information about the fiducial model assumed and mock datasets:

(i) We used WMAP - 5th year cosmology as our fiducial model with $[\Omega_m, \Omega_b, h, n_s, \sigma_8, w_0, w_a]$ as $[0.279, 0.0462, 0.701, 0.96, 0.817, -1.0, 0.0]$. We assume the equation of state of dark-energy is redshift dependent as (Chevallier & Polarski 2001; Linder 2003),

$$w(a) = w_0 + (1 - a)w_a \tag{32}$$

where, $a = 1/(1 + z)$ is the scale factor at redshift z.

(ii) We used three redshift bins to do the tomographic analysis with boundaries $[0.01, 0.8, 1.5, 4.0]$. So we calculated a total

of six spectra - three auto-spectra between bins 1-1, 2-2 and 3-3 and three cross-spectra between bins 1-2, 1-3 and 2-3.

(iii) We perform the likelihood analysis for different ℓ_{max} with $\ell_{min} = 10$ and 100 equally spaced logarithmic bins. So the bin sizes for the likelihood analysis with different ℓ_{max} are different.

(iv) We assumed that the mean redshift of the source distribution to be nearly 1.0 which gives approximately 50 galaxies per arc min^2 and $f_{\rm sky} = 0.55$ which resembles Euclid like survey.

(v) For the baryonic model, we used the realistic AGN feedback model $M_{\rm crit} = 10^{13}h^{-1}M_{\odot}$ as the fiducial value for total nine ℓ_{max} (1000, 2000, 3000, 4000, 5000, 6000, 8000, 10000, 20000). We also performed one case with more optimistic model $M_{\rm crit} = 10^{12}h^{-1}M_{\odot}$ for $\ell_{max} = 10000$. So there are ten cases in total.

(vi) We used our fiducial model stated above to generate shear power spectrum C_ℓ for these ten cases and perturbed all C_ℓ with normally distributed multi-variate random numbers drawn from a distribution with mean C_ℓ and the corresponding covariance matrix. These C_ℓ are catalogued and constitute the mock data sets. So, there are total ten mock data sets. In figure 4 we show the mock datasets up to $\ell_{max} = 20000$ for the six spectra and the best fits (which will be discussed in section 6.2).

For each bin combination (1-1,1-2 etc), the length of the data vector (ℓ or C_ℓ) is 100. Therefore, the total number of data points in each data set is 600. However, the two cross-spectra, 1-2 and 1-3, are highly correlated which actually leads us to have only 5 degree of freedom for each ℓ. Therefore, the total number of degree of freedom in each data set is about 492 (500 - 8 free parameters). Hence, the best fit to each dataset can have a χ^2 in the range $492 \pm \sqrt{(2 \times 492)}$ which is between 470 and 514.

In figure 3 (bottom-right panel), we show the boost for the unperturbed (without random noise) mock datasets up to very high ℓ_{max} with the corresponding DMO model. In all six curves of this figure, we kept $M_{\rm crit} = 10^{13}M_{\odot}$. The auto-spectra in the first bin (1,1), starts deviating (more than 1%) from the DMO model at about $\ell = 300$ whereas the auto-spectra of the third bin (3,3) starts showing deviation at nearly $\ell = 800$. All other auto-spectra and cross-spectra are between these two extremes. This behaviour is justified by looking at the same figure in upper-right panel, which shows the redshift evolution of the correction for the same $M_{\rm crit}$. It can be seen that at higher redshifts, the BAR matter power spectrum starts to deviate from DMO at smaller scales but also induces a larger dip at intermediate scales due to gas expulsion. This behaviour can be seen in the bottom-right panel. The C_ℓ in the lower redshift bin (1-1) starts deviating from DMO at larger scales as compared to the higher redshift bin (3-3), but the maximum dip in the two cases can be seen in the higher redshift bin (3-3). If we compare this to the bottom-left panel of the same figure, one can notice that the baryonic correction becomes even important when binning in redshift rather than using one big redshift bin. This provides additional constraints on $M_{\rm crit}$ while performing the analysis in tomographic bins compared to poorer constraints when only one bin is used.

5 LIKELIHOOD ANALYSIS AND COSMOLOGICAL IMPLICATIONS

We performed a likelihood analysis using MCMC to explore the cosmological parameter space for nine different ℓ_{max} (1000, 2000, 3000, 4000, 5000, 6000, 8000, 10000, 20000) using $M_{\rm crit} = 10^{13}h^{-1}M_{\odot}$, which is our most realistic model, and for $\ell_{max} = 10000$ using $M_{\rm crit} = 10^{12}h^{-1}M_{\odot}$ which is our optimistic model.

We run MCMC on the ten mock datasets obtained adopting both the DMO and BAR models, therefore we run a total of 20 MCMC. Each MCMC is performed using the publicly available code COSMOMC (Lewis & Bridle 2002), with 16 chains in each case. So total 320 CPUs are used for nearly 10 days to reach the desired convergence. The whole analysis required about 76800 hours.

We demonstrate the results of the MCMC and the interpretation in the following two sections, targeting particularly the precision and accuracy in predicting the cosmological parameters.

5.1 Precision in cosmology

Future experiments, like Euclid, are expected to provide very tight constraints on cosmological parameters. Here we show the constraints expected from using the weak lensing shear power spectrum as a function ℓ_{max}. Figure 5 shows the relative variance of four cosmological parameters and one baryonic parameter using both models, BAR (solid curves) and DMO (dashed curves). The matter density of the Universe (Ω_m) and the amplitude of fluctuations (σ_8) are the most constrained parameters, however, other parameters like the equation-of-state of dark-energy today (w_0) are relatively less constrained. The overall behaviour of all parameters is the same, weak constraints for small ℓ_{max}, better constraints with increasing ℓ_{max} and a flattening beyond $\ell_{max} \sim 8000$. The constraints derived from the BAR model are relatively weaker than the constraints derived from DMO model, which is the consequence of the extra parameter, $M_{\rm crit}$.

The normalized matter density of the Universe Ω_m can already be determined up to 5% at ℓ_{max}=1000 which improves as good as 2-3% at ℓ_{max}=8000 whereas the amplitude of fluctuations σ_8 can be determined much better at corresponding scales. At ℓ_{max}=1000, σ_8 can be known up to 3% and these constraints improves better than 1% at ℓ_{max}=8000. After ℓ_{max}=8000, the variance of both the parameters remains the same and no further constraints can be drawn by going up to lower scales or higher ℓ_{max}. There is a certain degeneracy in these two parameters which can be seen in figure 6 upper-left panel, where different colours represent different ℓ_{max}.

The constraints on the two parameters describing the redshift evolution of the equation of state of dark energy, w_0, w_a can also be improved with this kind of experiments. At ℓ_{max}=1000 w_0 can only be determined as good as 12%, whereas for ℓ_{max}=8000 it can be constrained up to 6-7% and with the same precision for higher ℓ_{max}. However, the constraints on w_a are much weaker. The absolute error on w_0 is nearly 0.35 for ℓ_{max}=1000, $\sim$0.18 for $\ell_{max} = 8000$ and the same afterwards.

The flattening of the relative errors of the parame-

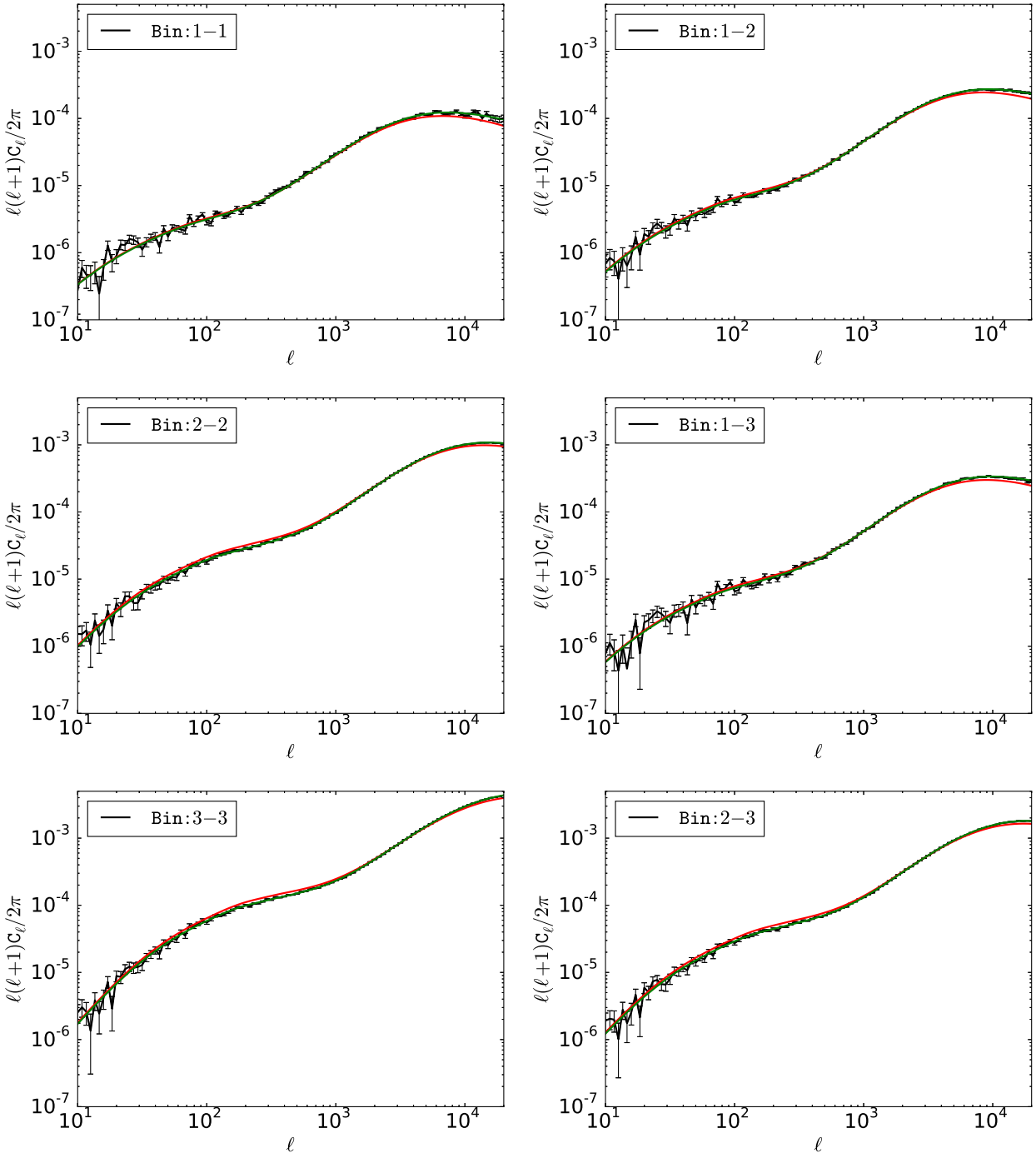

Figure 4. Mock datasets (including random noise) for $\ell_{max} = 20000$ in all six spectra (in black). The left column shows the three auto-spectra and the right column shows the three cross-spectra. Solid lines show the best fit for the DMO (red) and BAR (green) models.

ters indicates that there is no gain in precision of cosmological parameters estimation after a certain threshold $\ell_{max} \sim 8000$. In practice, an experiment like Euclid may provide us with very high quality data to even resolve and measure the shear power spectrum at ℓ_{max} as high as 10^5, but our analysis shows that the constraints becomes constant after $\ell_{max} \sim 8000$ and no further improvement can be achieved.

This forecast suggests that by measuring C_ℓs up to $\ell_{max} \sim 8000$, one can constrain Ω_m to about 2% precision and σ_8 to about 0.5% precision without any loss of information from high ℓs and including baryonic physics. However, w_0 can only be constraints up to 6-7% with some informa-

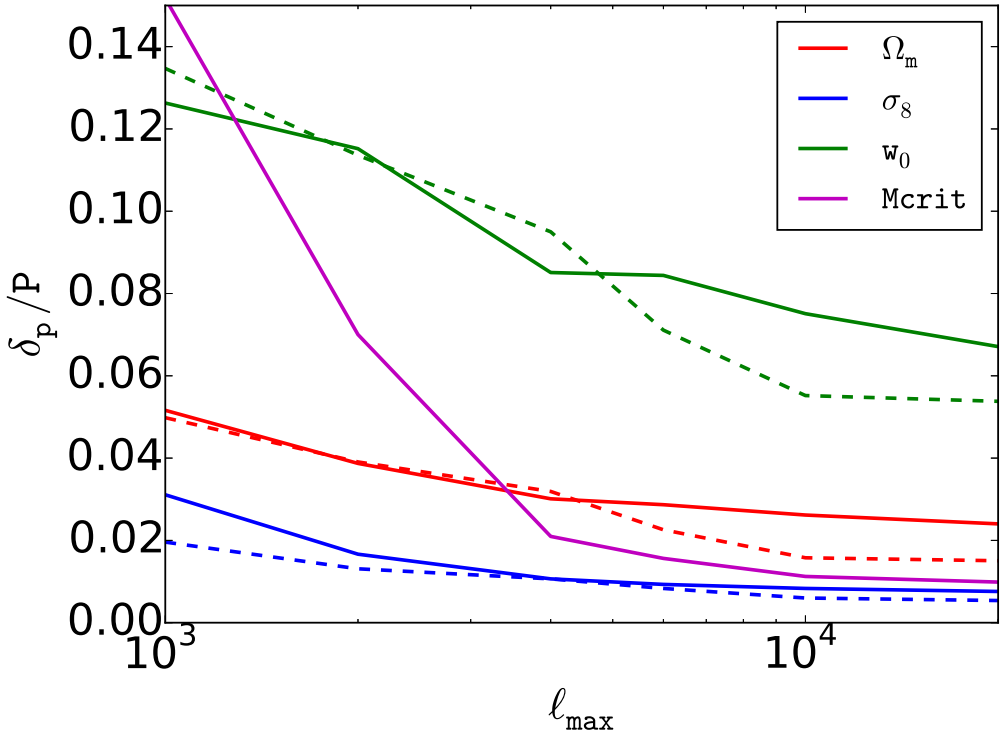

Figure 5. Relative 1σ errors on different cosmological parameters as a function of ℓ_{max} for $M_{\rm crit} = 10^{13}h^{-1}M_\odot$. Solid lines are for the BAR model and dashed curves are for the DMO model. Horizontal black dashed lines mark the ±1 and vertical black dashed lines shows important scales.

tion about w_a, the time derivative of the equation of state of dark energy.

5.2 Accuracy in cosmology

When precision cosmology is the goal, one should also take into account the ability to recover the cosmological parameter accurately. If there are systematic errors in the model, one can still derive very tight constraints from the wrong model, but the recovered parameters will be wrong or biased as compared to the true values. In this section we will present the results from our analysis of the bias in the cosmological parameters due to the lack of baryonic physics in DMO models and we will assess if these biases are significant. We define bias as the difference between the mean value of the parameter in MCMC and its fiducial or true value.

Figure 6 shows the 1σ error ellipses of cosmological parameters when the model is BAR (solid curves) and DMO (dashed curves). For small ℓ_{max}, the two models are indistinguishable, a consequence of the fact that baryonic physics becomes more important only at smaller scales. But as we go higher and higher in ℓ_{max}, the target density of the DMO model shifts further from the true target density, however, the BAR model remains at the correct location. We find that for all BAR models this bias is smaller than the 1σ error of the parameter, however, the bias in the parameters obtained fitting for the DMO model increases with increasing ℓ_{max}.

Figure 7 shows the ratio of these biases and the 1σ error on the cosmological parameters as a function of ℓ_{max} for the two models, BAR (solid curves) and DMO (dashed curves). The bias never exceeds the 1σ error for the BAR models, however, it does for the DMO models only after $\ell_{max} \sim 4000$. This is again a consequence of the fact that baryonic physics is only important at smaller scales. This indicates that if we only perform our experiment up to ℓ_{max}=4000, no baryonic physics needs to be taken into account, however, if one is interested in $\ell_{max} > 4000$ baryonic physics becomes very important. After ℓ_{max}=4000 the bias increases with ℓ_{max} and goes as big as 10σ at ℓ_{max}=10000 and remain flat after that. We see in the previous section that constraints on cosmological parameter can still be improved up to ℓ_{max}=8000, but considering the wrong model, DMO, the cosmological parameters will be 5-10σ away from the true values. So, in order to gain the best constraints on cosmology, baryonic physics must be taken into account.

5.3 An optimistic model

We analysed the $\ell_{max} = 10000$ case for our optimistic model with $M_{\rm crit} = 10^{12}h^{-1}M_\odot$. As in our previous analysis, we performed two MCMC in this case too, fitting for the BAR model and for the DMO model. Figure 6 (bottom row) shows the 1σ error ellipses of cosmological parameters. In this case the bias in the cosmological parameters does not exceed the 1σ error and hence is not a very troubling case. This was expected, as for lower $M_{\rm crit}$, baryonic physics is less important even at comparatively small scales as com-

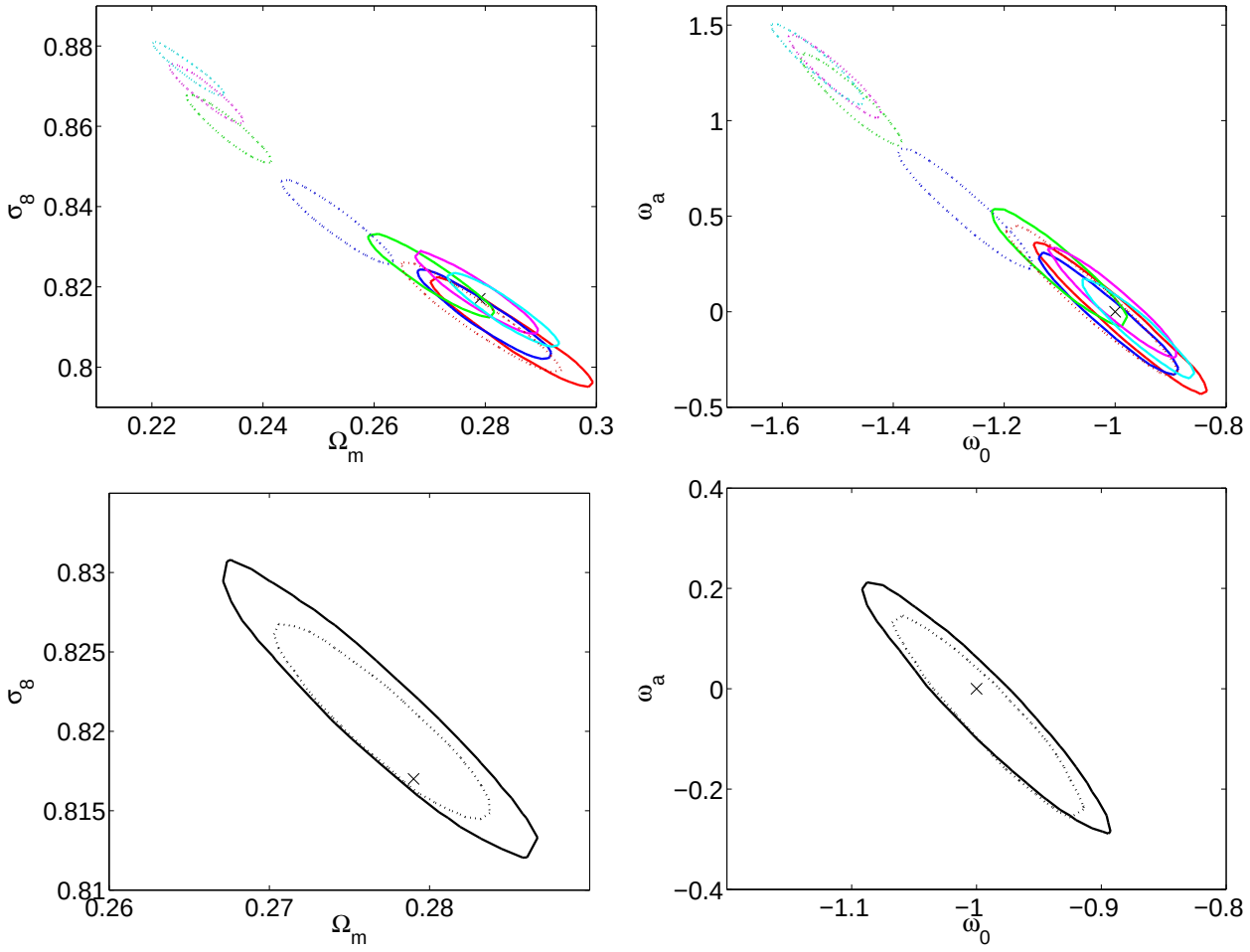

Figure 6. Top row: 1σ 2D error ellipses for different cosmological parameters using mock datasets with $M_{\rm crit} = 10^{13}h^{-1}M_{\odot}$ and different $\ell_{max} = 3000$ (red), 5000 (blue), 8000 (green), 10000 (magenta), 20000 (cyan). Bottom row: 1σ 2D error ellipses using mock datasets with $M_{\rm crit} = 10^{12}h^{-1}M_{\odot}$ for $\ell_{max} = 10000$. All solid curves are for the BAR model and dashed curves are for the DMO model.

pared to cases where $M_{\rm crit}$ is higher. For example if we look at figure 3 (bottom-left panel), we can see that for $M_{\rm crit} = 10^{12}h^{-1}M_{\odot}$, the deviation of C_ℓ from the DMO model is negligible at $\ell = 10000$. Hence, we actually expect smaller or no bias.

6 DISCUSSION AND CONCLUSIONS

In this work we first review the important theoretical framework necessary to calculate the matter power spectrum using the halo model and to compute the shear angular power spectrum in different redshift bins. We presented an analytic prescription to distribute baryons into two components – the intra-cluster plasma in hydrostatic equilibrium within the halo, and the BCG, which dominates the mass distribution in the centre of the halo, and whose properties are well measured using abundance matching techniques. We also take into account the adiabatic contraction of the dark matter particles due to the central condensation of baryons. We also compared these analytic density profiles to the simulations of Martizzi et al. (2014), both dark-matter-only and baryonic with AGN feedback, and found a remarkable agreement.

We model the shear power spectrum in the two models, BAR and DMO, and found that baryonic corrections are important after $k \sim 0.5\ h/Mpc$ in the matter power spectrum at redshift 0 for our most realistic AGN feedback model, which translates into $\ell \sim 800$ for the shear power spectrum in one big redshift bin. However, if binned in redshift space (lensing tomography), these corrections become larger in each bin and for each auto- and cross-correlation function. These baryonic corrections have one free parameter, $M_{\rm crit}$, which regulates AGN feedback, i.e., it controls how much gas will be inside the halo as a function of the halo mass. We believe the most realistic value of this parameter is near $10^{13}h^{-1}M_{\odot}$, which sets the most likely magnitude of baryonic corrections.

We perform the likelihood analysis using MCMC for total ten different datasets. Nine of them assume our realistic model for the AGN feedback with $M_{\rm crit} = 10^{13}h^{-1}M_{\odot}$ but different ℓ_{max}, and one assumes a less extreme (optimistic) model with $M_{\rm crit} = 10^{12}h^{-1}M_{\odot}$. For each mock dataset, we perform MCMC to fit for both models, BAR and DMO.

The main results of the likelihood analysis are summa-

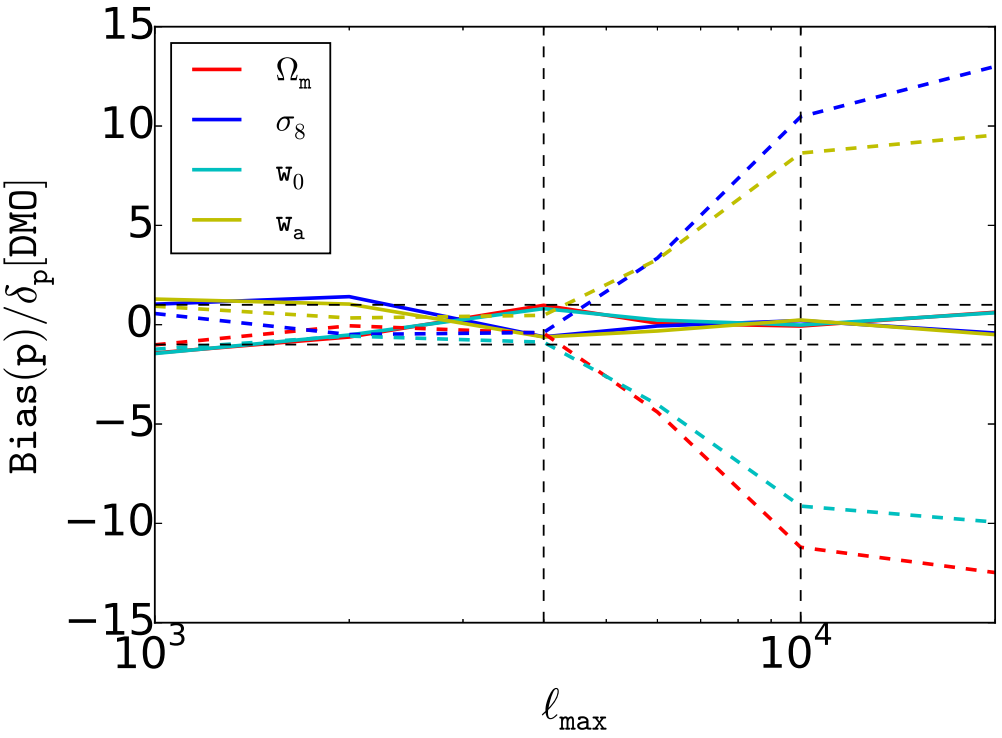

Figure 7. The ratio of bias and 1σ error of various cosmological parameters as a function of ℓ_{max} for $M_{\rm crit} = 10^{13}h^{-1}M_{\odot}$. Solid lines are for the BAR model and dashed curves are for the DMO model. Horizontal black dashed lines mark the ±1 and vertical black dashed lines shows important scales.

rized in figure 6, 5 and 7. The results are very interesting in two aspects: first, we found that the constraints on all cosmological parameters improve with increasing ℓ_{max}, but after $\ell_{max} \sim 8000$, the variance of each parameter becomes nearly constant. This indicates that even if we go to higher ℓ_{max} (or smaller scales), no additional constraints on the cosmological parameters can be gained. Second, if the wrong model, in this case DMO, is fitted to the data, after $\ell_{max} = 4000$ the mean recovered value of the parameters starts moving away from its true value. We refer to the difference between the true value and recovered mean value as bias in the cosmological parameter. The bias in the parameters becomes more than 1σ after $\ell_{max} = 5000$ and goes up to 10σ for $\ell_{max} = 10000$, remaining flat afterwards. So, there is a very interesting window from $\ell_{max} = 4000 - 8000$ which is useful for improving the constraints on cosmology, but if wrong model like DMO is chosen, the recovered cosmology can be highly biased from few to 10-σ.

6.1 Goodness of fit

In the previous sections we see that for $\ell_{max} < 4000$, there is no significant bias added to the determination of the cosmological parameters in our analysis, however, for $\ell_{max} > 5000$ the bias exceeds 1σ and keep increasing up to 10σ with increasing ℓ_{max}. The question here is: can we discard these biased models by looking at the goodness of fit? The answer to this question lies in figure 8 where we show the ratio between the best fit χ^2 in the DMO model and that in the BAR case as a function of ℓ_{max}. This ratio is as little as 5-10% up to $\ell_{max} \sim 5000$ but after that it only goes up to 25% at $\ell_{max} = 20000$ where bias is more than 10σ. Now, the reduced $\chi^2 = 1.25$ does not appear as such a bad fit for our cosmological measurements. So, by looking at the χ^2 only, it is not really possible to discard a model. The same conclusion can be drawn from figure 4, where we show the mock datasets of the six spectra (between different bins) for $\ell_{max} = 20000$. In this figure, we also show the two best fit from the DMO model (in red) and the BAR model (green). As we expect, the green curve is a better fit to the data than the red curve. But if the green curve is not present in this figure, the red curve does not appear to be a very bad fit. So, when deriving constraints on cosmology from this kind of experiments, one should be extremely careful about the possible magnitude of baryonic effects at small scales, because, although the results obtained with the wrong model may appear as a good fit, the corresponding bias can be in fact as high as many σ. Also, the recovered parameters from the wrong model (DMO) move away from the true value with increasing ℓ_{max}. This suggests a potential test for a given model, the cosmological parameter space should not move significantly when analysing up to different scales, the difference should only be seen in the variance of the parameters and not in its mean value.

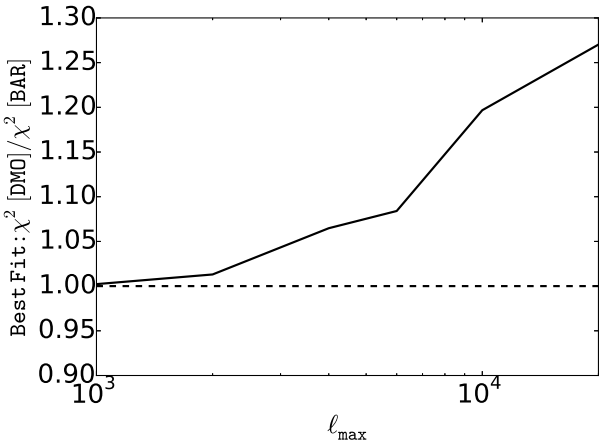

Figure 8. Showing the ratio of the best fit χ^2 in DMO model and BAR model at different ℓ_{max}.

6.2 $M_{\rm crit}$ parameter

The only free parameter in our BAR model, $M_{\rm crit}$, regulates the amount of gas inside the halo as a function of halo mass. We explore the consequences of what we believe to be a realistic model ($M_{\rm crit} = 10^{13}h^{-1}M_{\odot}$) in details considering nine different ℓ_{max}. At $\ell_{max} = 1000$, there is hardly any constraint drawn from weak lensing on this parameter, but as we increase ℓ_{max}, baryonic physics become more and more important and thus constraints can be put on $M_{\rm crit}$. In fact, the constraints on this parameter increase rapidly from 15% at $\ell_{max} = 1000$ to 1-2% at $\ell_{max} = 4000$. After this, no significant improvement on the constraints can be gained on this parameter. The variance of $M_{\rm crit}$ becomes constant after nearly $\ell_{max} = 8000$, which is what happens for the other cosmological parameters. So, with this kind of weak lensing experiment, $M_{\rm crit}$ (or $\log(M_{\rm crit})$) could be constrained up to 1-2%, which is quite impressive.

6.3 Non-Gaussian covariance vs baryonic corrections

Being able to extract cosmological information from clustering data down to a few percent accuracy can be considered very optimistic. It can be jeopardized by many unresolved issues. The two most important issues are (i) baryonic physics at small scales, and (ii) non-Gaussian effects in the covariance matrix of the power spectrum. These two issues can be quantified in projected weak-lensing statistics, like the shear power spectrum. In this work, we primarily talk about the effect of baryonic physics at small scales on the shear power spectrum and its cosmological implications. However, we ignore the effect of non-Gaussianity (NG) on the covariance matrix.

The NG contribution to the covariance becomes more important at small scales, like baryonic physics (Takada & Jain 2009; Takada & Hu 2013). Now the question is, which one is more important to deal with and which one appears first when going towards smaller scales? This question does not have a very straightforward answer. Ignoring both of these contributions may result in highly biased cosmological parameters estimations.

Yoo & Seljak (2012) (figure 9, right panel) shows the constraints on the amplitude of fluctuations (σ_8) as we go to smaller scales. If one considers only Gaussian errors, the constraints continue to improve until the instrumental shot noise kicks in. However, NG contribution are likely to dominate over Gaussian errors after $\ell = 700$. But we cannot directly compare to this plots as the constraints depend on many other details. We can still compare the ratios of the NG and Gaussian contributions. At $\ell_{max} = 10000$, the NG covariance is six times the Gaussian covariance. On the other hand, in figure 7 the bias in cosmology becomes close to 10σ for σ_8 at $\ell = 10000$. This means that the NG corrections are sub-dominant than the baryonic effects. However, our analogy is very hand-wavy and requires further study.

6.4 The ideal configuration

We explore the baryonic effects on the cosmological parameter estimation and found big bias in cosmological parameters if the analysis include $\ell > 4000$. After this limit, the cosmological parameters start to become biased and mislead the constraints. However, the constraints keep improving up to $\ell = 10000$. So the question arises, what is the ideal configuration to perform weak-lensing power spectrum analysis to put useful constraints on cosmology with Euclid-like surveys?

We explore this answer in our analysis and stated our results in the previous sections. To summarize, the ideal configuration is to go as high as $\ell = 8000$, including baryonic physics and marginalize over the baryonic parameters, in our case $M_{\rm crit}$. In this configuration, one can find unbiased estimates of the cosmological parameters. Having the unbiased estimates, we can also constrain the cosmological parameter space with much better accuracy than before. In this configuration, Ω_m and σ_8 can be estimated with nearly 2% and 0.5% respectively. The variance of the two parameters defining the redshift evolution of the equation of state of dark energy, w_0 and w_a are 0.07 and 0.15 respectively. Along with cosmological parameters, the baryonic parameter $M_{\rm crit}$ can also be estimated to very high accuracy, as good as 1-2%.

When dealing with real clustering datasets, we are also able to use independent constraints on the baryonic parameters, such as abundance matching data and/or X-ray data on individual halos, providing a solid understanding of the overall signal and the underlying baryonic effects.

7 ACKNOWLEDGEMENT

I.M. would like to thank Prasenjit Saha, Uros Seljak and Ravi Sheth for useful discussions about the topic and their suggestions.

REFERENCES

Albrecht A. et al., 2006, ArXiv Astrophysics e-prints
Amendola L. et al., 2013, Living Reviews in Relativity, 16, 6
Aubert D., Pichon C., Colombi S., 2004, MNRAS, 352, 376

Avila-Reese V., Firmani C., Klypin A., Kravtsov A. V., 1999, MNRAS, 310, 527
Bartelmann M., Schneider P., 2001, PhysicsReports, 340, 291
Bertschinger E., 2001, ApJS, 137, 1
Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Chevallier M., Polarski D., 2001, International Journal of Modern Physics D, 10, 213
Cimatti A. et al., 2009, Experimental Astronomy, 23, 39
Cooray A., Sheth R., 2002, PhysicsReports, 372, 1
Cui W., Borgani S., Dolag K., Murante G., Tornatore L., 2012, MNRAS, 423, 2279
Cusworth S. J., Kay S. T., Battye R. A., Thomas P. A., 2014, MNRAS, 439, 2485
Dehnen W., 2005, MNRAS, 360, 892
Duffy A. R., Schaye J., Kay S. T., Dalla Vecchia C., 2008, MNRAS, 390, L64
Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Eisenstein D. J., Hu W., 1999, ApJ, 511, 5
Eke V. R., Navarro J. F., Steinmetz M., 2001, ApJ, 554, 114
Gao L., Navarro J. F., Cole S., Frenk C. S., White S. D. M., Springel V., Jenkins A., Neto A. F., 2008, MNRAS, 387, 536
Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, ApJ, 616, 16
Gonzalez A. H., Zaritsky D., Zabludoff A. I., 2007, ApJ, 666, 147
Guillet T., Teyssier R., Colombi S., 2010, MNRAS, 405, 525
Hoekstra H., Jain B., 2008, Annual Review of Nuclear and Particle Science, 58, 99
Hu W., 1999, ApJL, 522, L21
Huterer D., 2002, PRD, 65, 063001
Jing Y. P., 2000, ApJ, 535, 30
Jing Y. P., Zhang P., Lin W. P., Gao L., Springel V., 2006, ApJL, 640, L119
Kravtsov A., Vikhlinin A., Meshscheryakov A., 2014, ArXiv e-prints
Laureijs R. et al., 2011, ArXiv e-prints
Lewis A., Bridle S., 2002, PRD, 66, 103511
Linder E. V., 2003, Physical Review Letters, 90, 091301
Ma C., Fry J. N., 2000, ApJ, 543, 503
Macciò A. V., Dutton A. A., van den Bosch F. C., Moore B., Potter D., Stadel J., 2007, MNRAS, 378, 55
Martizzi D., Mohammed I., Teyssier R., Moore B., 2014, MNRAS, 440, 2290
Martizzi D., Teyssier R., Moore B., 2013, MNRAS, 432, 1947
Martizzi D., Teyssier R., Moore B., Wentz T., 2012, MNRAS, 422, 3081
Mashchenko S., Couchman H. M. P., Wadsley J., 2006, Nature, 442, 539
McClelland J., Silk J., 1977, ApJ, 217, 331
Mellier Y., 1999, ARAA, 37, 127
Mohammed I., Seljak U., 2014, ArXiv e-prints
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Muñoz-Cuartas J. C., Macciò A. V., Gottlöber S., Dutton A. A., 2011, MNRAS, 411, 584
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Neto A. F. et al., 2007, MNRAS, 381, 1450
Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144
Pontzen A., Governato F., 2012, MNRAS, 421, 3464
Read J. I., Gilmore G., 2005, MNRAS, 356, 107
Reddick R. M., Tinker J. L., Wechsler R. H., Lu Y., 2014, ApJ, 783, 118
Refregier A., 2009, Experimental Astronomy, 23, 17
Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, ApJ, 672, 19
Schaller M. et al., 2014, ArXiv e-prints
Schneider P., 2006, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro, Meylan G., Jetzer P., North P., Schneider P., Kochanek C. S., Wambsganss J., eds., pp. 269–451
Seljak U., 2000, MNRAS, 318, 203
Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020
Stanek R., Rudd D., Evrard A. E., 2009, MNRAS, 394, L11
Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
Takada M., Hu W., 2013, PRD, 87, 123504
Takada M., Jain B., 2004, MNRAS, 348, 897
Takada M., Jain B., 2009, MNRAS, 395, 2065
Teyssier R., 2002, A&A, 385, 337
Teyssier R., Moore B., Martizzi D., Dubois Y., Mayer L., 2011, MNRAS, 414, 195
Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, MNRAS, 429, 3068
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
van Daalen M. P., Schaye J., Booth C. M., Dalla Vecchia C., 2011, MNRAS, 415, 3649
White M., 2004, Astroparticle Physics, 22, 211
Yoo J., Seljak U., 2012, PRD, 86, 083504
Zhan H., Knox L., 2004, ApJL, 616, L75
Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2003, ApJL, 597, L9

5

Paper4: Lensing time delays as a substructure constraint

Strong gravitational lensing (SL) is a very powerful tool to reconstruct the mass distribution of the lens given the positions of the multiple images of the background source(s). Another observable of the SL phenomenon is the time delay between two images of the same source, due to the fact that two light rays from the same source, emitted at the same time, travel different paths (and so the distances), and reach the same observer. If the source is variable, like a quasar or a supernovae etc., it is possible to measure these time delays.

As described by equation 1.49, the total time delay between two images is proportional to the comoving distance from the observer to the lens, and therefore depends on the Hubble constant H_0. If the mass distribution of the lens is known, it is possible to infer the H_0 by measuring the time delays. However, this exercise is very sensitive to the mass distribution of the lens. Therefore, a very good understanding of the gravitational potential of the lens is needed in order to measure the H_0 using the time delays. This has been done by many authors in the past, where they derived necessary constraints on H_0.

In this paper, we quantified the additional constraints provided by the time delays on the mass distribution of the lens assuming a cosmological model (or H_0). We performed a principle component analysis (PCA) using two types of mass maps of the lensing cluster SDSS J1004+4112 - one reconstructed with time delays data (TD) and other without it (NTD). The main science driver of this paper is to identify and isolate those uncertainty modes that are present in NTD maps and are not in TD maps. By successfully identifying these modes, we concluded that in the lensing clusters where the steepness degeneracy is already broken by multiple background

sources at different redshifts, time delay information can be used to constrain the lopsidedness of the cluster core.

Role: Under the supervision of Dr. Prasenjit Saha, I started this project by accumulating lensing data for SDSS J1004 cluster. Along with the multiple images of three background sources, two time delays were also available for one of the background sources, which is a quasar. Employing this dataset, I reconstructed 30 mass maps using lensed images only, and 30 using lensed images plus two time delays. I also computed the average mass map with optimal resolution in both cases. The next step was to isolate the uncertainty modes that were present in NTD maps, and not in TD maps. Each mass map is a grid of pixels, and for each case total 30 maps were available. I computed the uncertainty/moment matrix for NTD models. I isolated its principle component, which was the eigen-vector corresponding to the largest eigen-value, that resembles the largest uncertainty mode present in NTD models, and compared its effect to the reference mass model that we chose to be the average TD mass map.

This paper has been published in *Publication of Astronomical Society of Japan* (PASJ).
Arxiv: `http://arxiv.org/abs/1412.3464`

Publications of the Astronomical Society of Japan Advance Access published March 12, 2015

Publ. Astron. Soc. Japan (2015) 00 (0), 1–9
doi: 10.1093/pasj/psu155
Advance Access Publication Date: 2015 0

Lensing time delays as a substructure constraint: a case study with the cluster SDSS J1004+4112

Irshad Mohammed,[1,2,*] Prasenjit Saha,[1,2] and Jori Liesenborgs[3]

[1]Physik-Institut, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
[2]Institute for Computational Science, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
[3]Expertisecentrum voor Digitale Media, Universiteit Hasselt, Wetenschapspark 2, B-3590 Diepenbeek, Belgium

*E-mail: irshad@physik.uzh.ch

Received 2014 November 6; Accepted 2014 December 10

Abstract

Gravitational lensing time delays are well known to depend on cosmological parameters, but they also depend on the details of the mass distribution of the lens. It is usual to model the mass distribution and use time-delay observations to infer cosmological parameters, but it is naturally also possible to take the cosmological parameters as given and use time delays as constraints on the mass distribution. In this paper we develop a method of isolating what exactly those constraints are, using a principal-components analysis of ensembles of free-form mass models. We find that time delays provide tighter constraints on the distribution of matter in the very highly dense regions of the lensing clusters. We apply it to the cluster lens SDSS J1004+4112, whose rich lensing data include two time delays. We find, assuming a concordance cosmology, that the time delays constrain the central region of the cluster to be rounder and less lopsided than would be allowed by lensed images alone. This detailed information about the distribution of the matter is very useful for studying the dense regions of the galaxy clusters which are very difficult to study by direct measurements. A further time-delay measurement, which is expected, will make this system even more interesting.

Key words: galaxies: clusters: individual (SDSS J1004+4112) — gravitational lensing: strong

1 Introduction

While ΛCDM cosmology is a very successful framework, the underlying nature of both Λ (dark energy) and CDM (cold dark matter) remains unknown. The interaction of baryons with both of these is well understood during the linear-growth era of structures, but less so when clusters and galaxies start to form. *N*-body simulations of CDM give dark-matter distributions which roughly follow cuspy profiles with a characteristic radius, "well-known NFW (Navarro–Frenk–White) profiles" (Navarro et al. 1997); however, hydrodynamical simulations and other analytic studies show that in the presence of baryons, NFW is no longer a good fit in the innermost part of the haloes (see for example Schaller et al. 2014; Mohammed et al. 2014b). The distribution of matter in the innermost parts of the galaxy clusters is dominated by the baryonic component, particularly by BCG (brightest cluster galaxies)

and other elliptical galaxies. So, the distribution is different from that of dark matter and does not follow an NFW profile, which is a very good fit in the outskirts of the cluster. Due to the high potential well of galaxies, some dark matter contracts adiabatically, making its profile steeper at the centre (Blumenthal et al. 1986; Gnedin et al. 2004; Read & Gilmore 2005). Generally, the centres of the BCG host active nuclei (AGN), which push through feedback the gas near the centre of the halo to the outskirts (Dehnen 2005; Mashchenko et al. 2006; Martizzi et al. 2013; Teyssier et al. 2013). For low-mass halos AGN feedback can push all the gas outside the halo whereas for high-mass halos AGN feedback is not so strong. At the very transition from big groups and galaxy clusters, the AGN feedback is strong enough to push some gas outside the halo but not all. These processes make the centre of the halo very dynamic and redistribute the matter near the centres of galaxy clusters. It is difficult to resolve the structures in those high-density regions by direct observations, however, strong gravitational lensing (SL) is capable of resolving those scales (Hammer 1991; AbdelSalam et al. 1998a; Kneib 2002; Halkola et al. 2008; Liesenborgs et al. 2009; Mohammed et al. 2014a; Sharon et al. 2014). The precise identification of the multiply imaged background galaxies/quasars at different redshifts can make SL very powerful. However, there are still degeneracies that preclude strong constraints on the central regions of galaxy clusters. In this paper we suggest that lensing time delays may provide additional information on the central substructure.

The idea of measuring time delays in multiply imaged lensed systems was discussed theoretically long before any was discovered. Refsdal (1964, 1966) notably brought some remarkable insights, which we may summarise as follows. First, the time delays due to a lens of mass M is of the order of GM/c^3, hence weeks to years for galaxy and cluster masses, which is conveniently human-scaled. Secondly, whereas the image data on a lens are all angular quantities and hence dimensionless, a time delay introduces a dimension, which in fact is proportional to the Hubble time. Thirdly, since lensing depends on the ratio of source–lens and observer–lens distances, and these distances depend on the cosmological model, time delays coming from different redshifts can potentially measure the cosmological parameters.

However, lensing time delays also depend on the mass distribution of the lens, and this introduces uncertainty. To measure cosmological parameters one needs a strong prior (Sereno & Paraficz 2014), especially if only a single lens is used (Suyu et al. 2014). If only the Hubble time is sought, while the Ω parameters are assumed, an ensemble of lenses gives better-constrained results (Saha et al. 2006; Oguri 2007; Coles 2008; Paraficz & Hjorth 2010; Rathna Kumar et al. 2014), but still not as precise (so far) as from the cosmic microwave background.

In this paper, we reverse the traditional process, and take the cosmological parameters as given. Now, in lenses with sources at only one redshift and given cosmological parameters, a time delay breaks the steepness degeneracy. But in situations like SDSS J1004+4112, where the steepness is already partly constrained by other lensing observables, time delays provide information about the shape of the mass distribution, particularly very close to the centre, which is very difficult to probe with direct observations. These central regions are also very difficult to probe with weak lensing or flexion data. We develop a method of quantifying what time-delay measurements tell us about a lensing mass distribution. We then apply the method to SDSS J1004+4112, which has two measured time delays. A further time delay is expected, so the results and interpretation are preliminary. Nonetheless, they provide an insight into what may be possible.

2 The cluster SDSS J1004+4112

The cluster SDSS J1004+4112 at redshift 0.68 has three strongly lensed systems. At redshift 1.74, there is a quasar (Q) lensed into five images (Q1–Q5) (Inada et al. 2003, 2005). Further, there is a galaxy (A) at redshift 3.332 lensed into five images, and another galaxy (B) at redshift 2.74 is lensed into two images (Sharon et al. 2005) (see figure 1).

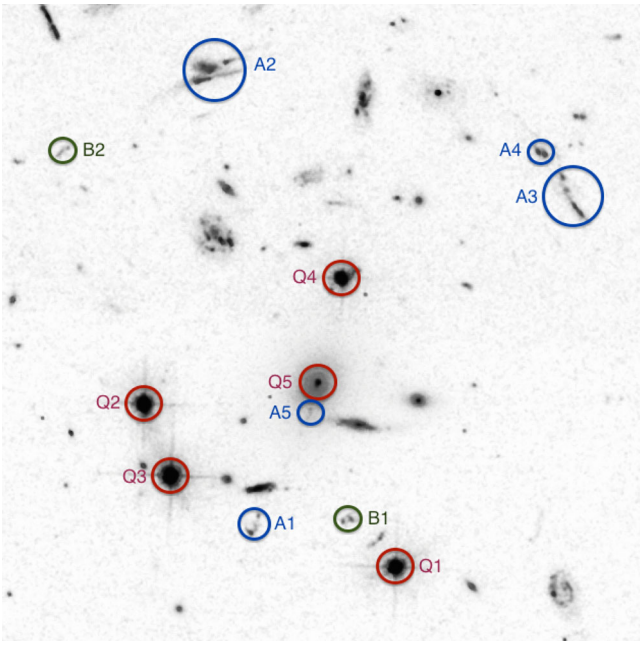

Fig. 1. Multiple images of the quasar, labelled Q1 to Q5 in order of time delay (measured or expected). Also showing galaxy A (A1-A5) and B (B1-B2) images. (Color online)

Still another candidate for lensed system (C) is known, but not yet confirmed spectroscopically.

The Q system is natural for time-delay measurements. Two of the four possible time delays (between Q1 and Q2, and between Q1 and Q3) of the quasar images have been measured (Fohlmeister et al. 2007, 2008) and a third is expected. The image separation is large (up to 14″), and since time delays are proportional to the square of the image separation, the time delays are much longer than with galaxy lenses. Image Q3 lags behind the nearby Q2 by 40 d and lags behind Q1 by 821 d. The cluster gas has also been observed in X-rays (Ota et al. 2006). As data have been accumulated, many different models have been published (Oguri et al. 2004; Williams & Saha 2004; Kawano & Oguri 2006; Saha et al. 2007; Inada et al. 2008; Liesenborgs et al. 2009; Oguri 2010).

3 Time delays in lensing

Let us briefly recall the physics of time delays. This is conveniently done using Fermat's principle as applied to gravitational lensing (Blandford & Narayan 1986). Consider a sky-projected density $\Sigma(\boldsymbol{x})$ at redshift $z_{\rm L}$. Here $\boldsymbol{x}$ represents the physical (not comoving) coordinates perpendicular to the line of sight. The gravitational time delay due to this mass will be

$$\nabla^2 t_{\rm grav} = -(1+z_{\rm L})\frac{8\pi G}{c^3}\Sigma(\boldsymbol{x})\,. \tag{1}$$

Now consider a source at $\boldsymbol{x}_{\rm s}$. A light ray coming from this source, being deflected at the lens so that it heads to the observer, will have an additional geometrical time delay of

$$t_{\rm geom} = \frac{(1+z_{\rm L})}{2c}\frac{d_{\rm S}}{d_{\rm L}d_{\rm LS}}|\boldsymbol{x}-\boldsymbol{x}_{\rm s}|^2\,, \tag{2}$$

where the $d_{\rm L}$ is the angular-diameter distance to the lens, and $d_{\rm S}$ and $d_{\rm LS}$ are angular-diameter distances from the observer to the source and from the lens to the source, respectively. In a flat cosmology

$$d_{z_1,z_2} = \frac{c/H_0}{1+z_2}\int_{z_1}^{z_2}\frac{dz}{\sqrt{\Omega_{\rm m}(1+z)^3+\Omega_\Lambda}} \tag{3}$$

gives the various angular-diameter distances.

The total time delay is then

$$t(\boldsymbol{x},\boldsymbol{x}_{\rm s}) = t_{\rm geom}(\boldsymbol{x},\boldsymbol{x}_{\rm s}) + t_{\rm grav}(\boldsymbol{x})\,. \tag{4}$$

Images will form where $t(\boldsymbol{x},\boldsymbol{x}_{\rm s})$ has a minimum, saddle point, or maximum. The tensor magnification is the inverse matrix of second derivatives of $t(\boldsymbol{x})$. Note that the dependence on $\boldsymbol{x}_{\rm s}$ has been differentiated out. Flexion consists of the derivatives of the tensor magnification, hence the third derivatives of the time.

Lens modelling consists of reconstructing $\Sigma(\boldsymbol{x})$ and $\boldsymbol{x}_{\rm s}$. For a quasar source, $\boldsymbol{x}_{\rm s}$ is a single point, and for an extended source a superposition of source points must be considered. The earliest detailed lens models (Young et al. 1981) already noted the non-uniqueness of lens models. Falco, Gorenstein, and Shapiro (1985) quantified the most important issue of these, now known as the steepness degeneracy or the mass-sheet degeneracy; steeper mass profiles give longer time delays, while leaving image positions and shapes the same. As modellers explored models further, it turned out that the shape of the mass distribution also affects time delays (Saha & Williams 1997, 2006; Zhao & Qin 2003). Degeneracies have also been studied theoretically (Schneider & Sluse 2013). Having sources at multiple different redshifts (high-redshift contrast) tends to suppress degeneracies (AbdelSalam et al. 1998b; Saha & Read 2009) but does not eliminate them completely (Liesenborgs et al. 2008; Liesenborgs & De Rijcke 2012).

It is useful to break down lensing time delays into three factors: lens substructure, lens size, and cosmology. The time delay between the innermost and outermost images can be written as

$$\Delta t_{\rm in,out} = f_{\rm lens}\frac{A_{\rm lens}}{A_{\rm sky}}\frac{d_{\rm L}d_{\rm S}}{d_{\rm LS}}, \quad A_{\rm lens} = \frac{\pi}{4}(\theta_{\rm in}+\theta_{\rm out})^2. \tag{5}$$

Cosmology enters through the distance factors, while $A_{\rm lens}$ is the size of the lens on the sky, and is fixed by that the astrometry. With these factors fixed, $f_{\rm lens}$ is the remaining dependence on substructure. Typical values are 2–6 for systems with 2 + 1 images, and 0.5–2 for systems with 4 + 1 images (Saha et al. 2006). That is to say, substructure is very important for time delays. This dependence is undesirable when estimating cosmological parameters, but it is welcome for inferring substructure.

4 Isolating the time-delay modes

In this section we produce a form of principal components analysis (PCA) to isolate the information that time delays provide on the mass distribution, assuming the cosmological parameters are known. A somewhat related technique, for lensing clusters with multiple source redshifts but not necessarily including time delays, is developed by Lubini et al. (2014).

We reconstruct the mass distribution in two ways: first, including the measured time delays (say TD models), and secondly, with no time-delay information (say NTD models). For each of TD and NTD, we reconstructed an ensemble of mass maps (30 in number). Each mass map is a grid map (size 74 × 74).

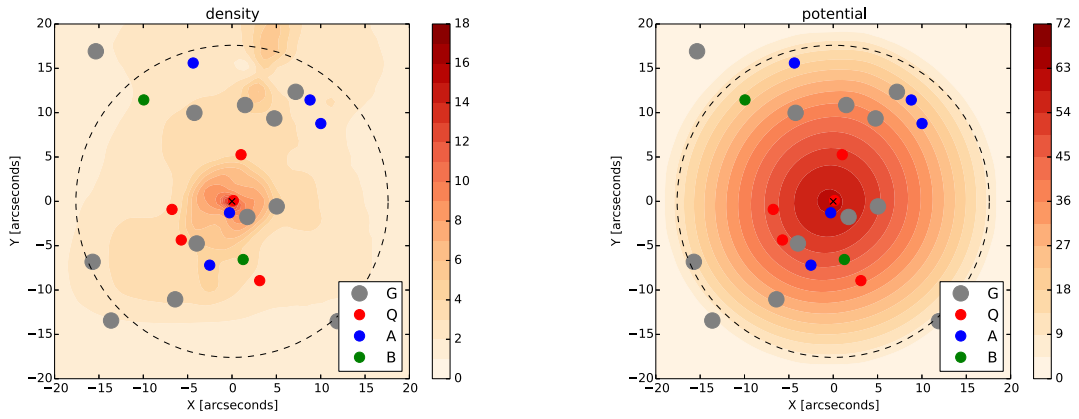

Fig. 2. The reference model, [equation (6)], using all the data constraints. Density (left-hand panel) means projected density in kg m^{-2} and potential (right-hand panel) is $t_{\rm grav}$ in years. Lensed images and cluster galaxies are marked.

We now wish to find the variation present in the NTD ensemble but not in the TD ensemble. This will provide information on substructure possibilities left by image data (only) but ruled out by time delays. Let X^i_n denote the projected density of the ith TD model (i going from 1 to 30 in this paper) at the nth grid point (of $74^2 = 5476$ grid points). Similarly use Y^i_n for the NTD mass maps. Next, we choose a reference Z_n,

$$Z_n = \langle X^i_n \rangle , \tag{6}$$

which is the ensemble average of the TD maps. Then,

$$\Delta X^i_n = X^i_n - Z_n \tag{7}$$

is the ensemble variation from the reference. Then we introduce a moment matrix

$$\mathbf{M}_{mn}(X) = \langle \Delta X^i_m \, \Delta X^i_n \rangle , \tag{8}$$

where the average is again over the TD maps. $\mathbf{M}_{mn}(X)$ is just the covariance between pairs of grid points. The eigenvalues and eigenvectors of $\mathbf{M}_{mn}(X)$ describe how sets of grid points tend to vary together. Let us denote them by $\lambda_k(X)$ and $V^k_n(X)$, respectively; the superscript k denotes the kth eigenvalue and eigenvector ($k = 1$ having the largest eigenvalue). In practice, the first few eigenvalues dominate. The vector

$$Z_n \pm \sqrt{|\lambda_k(X)|}\, V^k_n(X), \tag{9}$$

displayed as a density map, shows the principal mode of variation of the mass model. (There is no sum over k.) These variation modes are, of course, orthogonal.

We then proceed to the NTD maps. Let these be Y^i_n and let

$$\Delta Y^i_n = Y^i_n - Z_n \tag{10}$$

be the variations with respect to the reference model. We now subtract the variation modes of TD mass maps

$$\Delta \bar{Y}^i_n = \Delta Y^i_n - \sum_k \left[\textstyle\sum_m \Delta Y^i_m V^k_m(X) \right] V^k_n(X), \tag{11}$$

leaving variations that are orthogonal to the TD variations. Using these, we build another moment matrix

$$\mathbf{M}_{mn}(Y) = \langle \Delta \bar{Y}^i_m \, \Delta \bar{Y}^i_n \rangle . \tag{12}$$

The eigenvalues and eigenvectors of $\mathbf{M}_{mn}(Y)$ contain the modes present in NTD map models but absent from the TD maps. These modes can be conveniently displayed as

$$Z_n \pm \sqrt{|\lambda_k(Y)|}\, V^k_n(Y). \tag{13}$$

We call these "variations ruled out by the time delays" and these are the main results of this paper. One sign of the $\pm$ must be chosen for definiteness, but it does not matter which one.

5 Application to SDSS J1004+4112

The mass distribution $\Sigma(\boldsymbol{x})$ can be reconstructed in different ways; Oguri (2010) uses a parametrized form, and Saha, Williams, and Ferreras (2007) built it out of mass tiles or pixels, while Liesenborgs et al. (2009) used an adaptive superposition of Plummer components. The latter, and in particular the GRALE code, is used in the present work.

5.1 Mass reconstruction using GRALE

As mentioned above, GRALE (Liesenborgs et al. 2006, 2007, 2009) makes free-form mass models for the lensing cluster, as a superposition of many Plummer lenses. Except for the redshift and general location, no information from the lens itself is used. The inversion input consists of

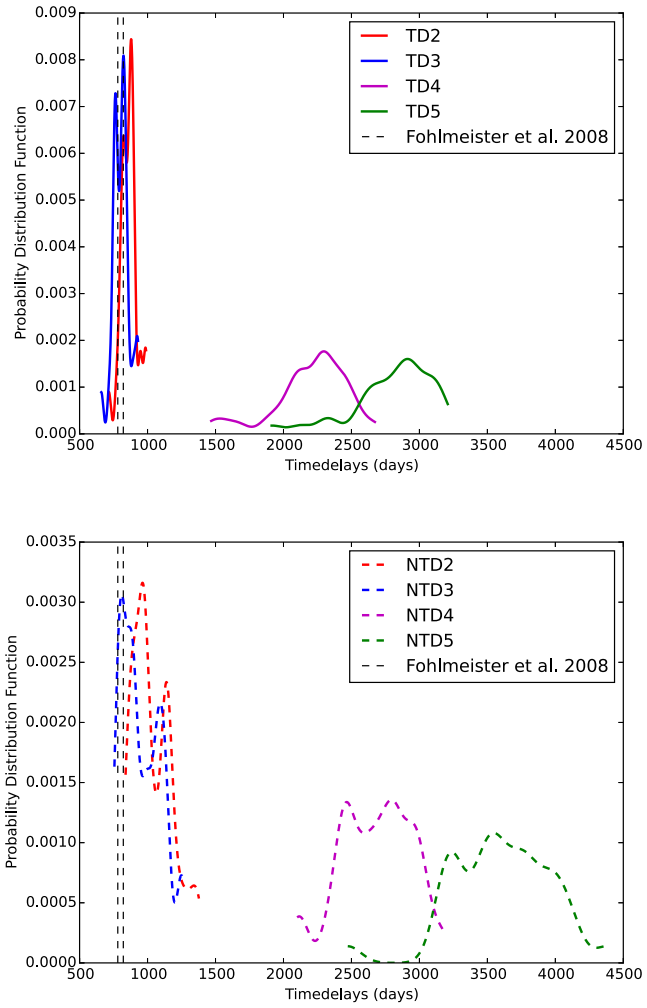

Fig. 3. Time delay estimates using TD models (upper) and NTD models (lower). TD2, TD3, TD4, and TD5 (or NTD2, NTD3, NTD4, and NTD5) are the time delays for images Q2, Q3, Q4, and Q5, respectively, with respect to Q1. Vertical dashed lines show the input time delays from Fohlmeister et al. (2008).

(1) lensed-image positions and source redshifts, (2) regions where additional images have not been identified but could be present, and (3) time delays, if any. Each of these is used to define a fitness measure of a mass distribution.

(i) For a given mass map, the input multiple images are ray-traced back to the source plane. The "overlap fitness" of the mass map expresses how well these back-projected images overlap. It is important to consider fractional overlap rather than simple source–plane distances in order to avoid favouring extreme magnification (tiny sources).

(ii) The back-projection may give rise to extra images. If they are not in regions specified by the modeller as allowed, they are spurious. These penalise the mass map through a "null fitness" measure. Through the null fitness, GRALE uses the non-occurrence of images at random locations as useful data.

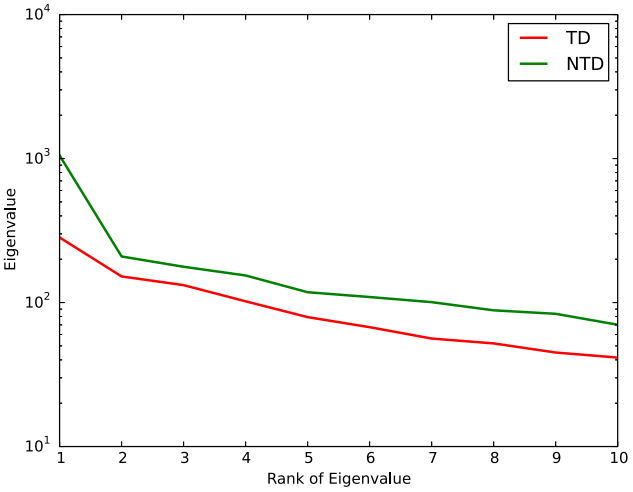

Fig. 4. Distribution of eigenvalues $\lambda_k(X)$ and $\lambda_k(Y)$, labelled as TD and NTD, respectively, ranked by absolute value.

(iii) The "time-delay fitness" measures how well the time delay in a mass distribution agrees with the observations.

GRALE uses a multiobjective genetic algorithm to find free-form mass maps which provide optimal fits to the data on the above criteria. If more Plummer-lens components are allowed, the fitness will tend to be better. Accordingly, we used a heuristic Occam's razor criterion to find a compromise between better fitness and more components (Mohammed et al. 2014a). This effect sets the resolution adaptively. No additional priors or regularization are used.

5.2 Mass models

Our mass maps for SDSS J1004+4112 were of two types, as follows.

(i) No time delay (or NTD) models. We used the three image systems (Q, A, and B; total 12 images), but gave no time-delay information in this case. GRALE finds an optimal solution, restricting extra images using the technique of null spaces, for this data set. We repeated the same procedure 30 times to generate a statistical ensemble of 30 models.

(ii) Time delay (or TD) models. For these we used the same data set, plus the two measured time delays: 40 d in Q2–Q3 and 821 d in Q1–Q3 (see figure 1). Following the same procedure, we made an ensemble of 30 mass maps. In figure 2 show the ensemble average of the TD models, which is used as the reference Z_n from section 4. The left-hand panel of figure 2 shows the projected density and the right-hand panel shows the potential in the colour code. This mass distribution is

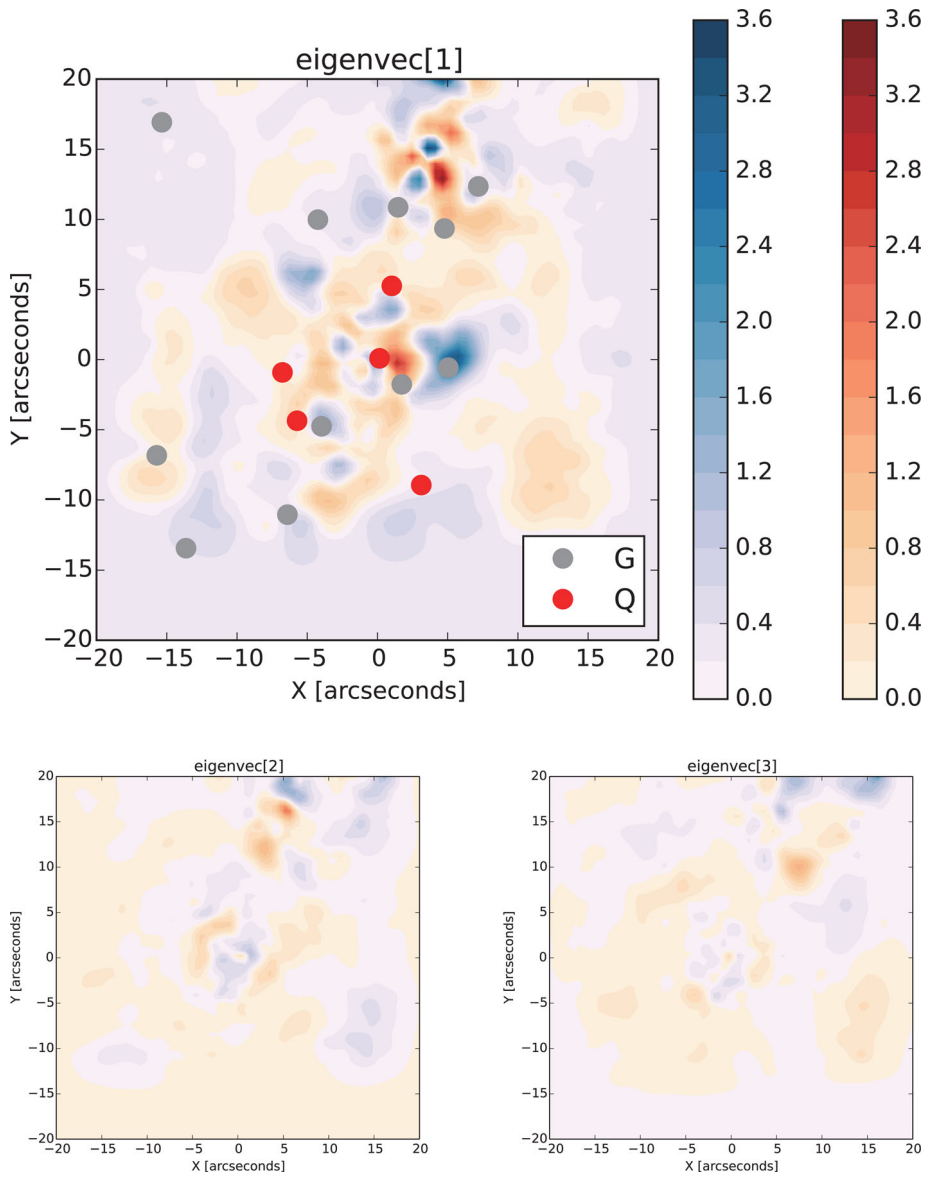

Fig. 5. The main/largest variation in surface density ruled out by time delays (upper panel). The red (right) and the blue (left) color bars represents positive and negative values (in kg m^{-2}), respectively. The two lower panels show the second and third largest variation modes ruled out by time delays.

slightly different from that of Liesenborgs et al. (2009), because the present work uses an adaptive resolution scheme, but these differences do not appear to be significant.

The method described in section 4 is then applied to the TD and NTD models. Tests in Mohammed et al. (2014a) indicate that GRALE ensembles of this size underestimate the actual uncertainties by a factor of two, but do explore the different uncertainties. Thus, the eigenvalues reported below are certainly underestimates, but the eigenvectors should be a good representation of the variation.

5.3 Results and interpretation

Figure 3 shows the distribution of time delays in the TD and NTD ensembles. As expected in the TD models, the first two time delays, which were also used as data inputs, show little variation (though larger than the observational uncertainties) while the other two show large variations. In the NTD models, all the time delays show large variations. Most of the TD models also favour lower time delays for Q4 and Q5, as compared to the NTD models estimate. Thus, existing time-delay measurements constrain future measurements to some degree, but not very tightly, indicating that future measurements will bring substantial new information.

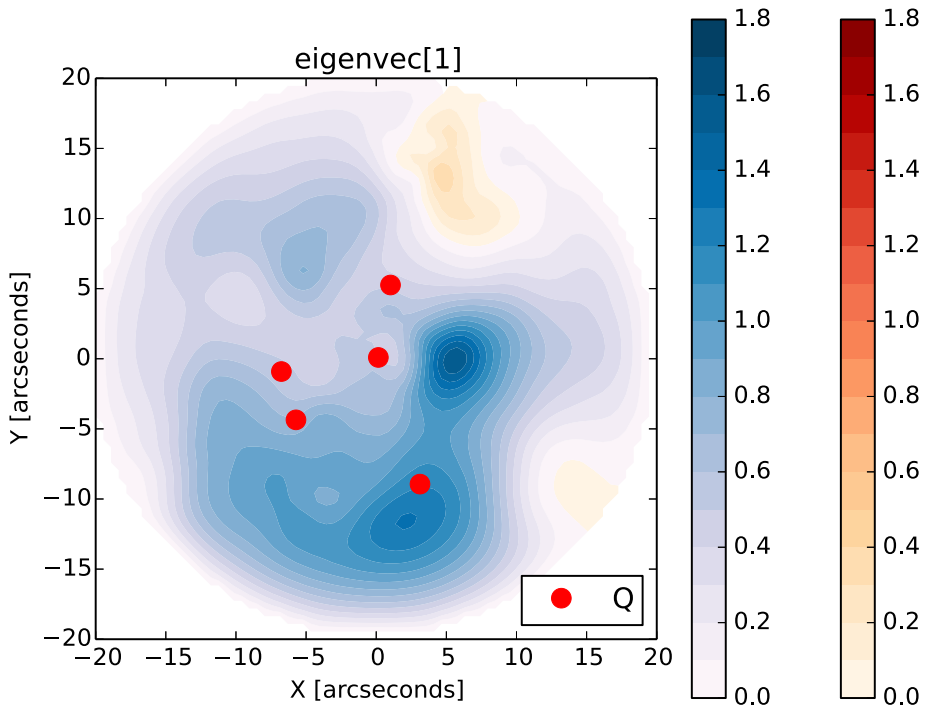

Fig. 6. The main/largest variation in lens potential ruled out by time delays. The red (right) and the blue (left) colour bars represent positive and negative values (in years), respectively.

Figure 4 shows the eigenvalues of the NTD and TD modes. The largest NTD mode evidently dominates, being a factor of five larger than the second largest mode. When time delays are included, it is also the case that one TD mode is much larger than the rest.

Figure 5 presents the main result of this paper. It illustrates the largest NTD variation mode that is absent in the TD ensemble [cf. equation (13)], in other words, the largest variations ruled out by the time-delay measurements. The upper panel of the figure shows the eigenvector corresponding to the largest eigenvalue. The two central blobs in the upper panel of figure 5 (blue and red, hence anti-correlated) allow the mass to shift the local mass peak towards or away from the cluster centre, and hence change the steepness of the mass profile. Without the TD constraints, the central peak of the mass distribution can be more lopsided, whereas the TD constraints force the central peak to be rounder. The rather large uncertainties (20%–40% in Σ) in the central region of the NTD models is reduced in the TD models. This local uncertainty of 20%–40% is, however, less than 1% of the total mass in the strong-lensing region. That is to say, time delays are constraining substructures that are only a percent of the total. The red and blue blobs are also in the vicinity of the two galaxies in the very central region. That may be coincidental, but it is worth remarking that the mass in the blobs is of the order of a galaxy-halo mass such as may be tidally stripped from a galaxy near the centre of a cluster. This suggests that time delays may give very sophisticated information about the variation of the mass near the environment of the galaxies in the central region, or in other words, about the substructures in a dense environment. Therefore, the third time delay, which is expected for the image Q4, could be very useful in extracting the substructure information of the cluster near its centre. Note that this interesting region is not accessible through the more well-known techniques of weak lensing and flexion.

The two lower panels of figure 5 show the second and third largest variation modes. These variation modes are weaker, also evident from figure 4, as all other eigenvalues after the first one are subdominant. The number of non-zero eigenvalues for NTD should equal the number of time-delay measurements (two in this case), but in practice further modes are present, and only gradually die away. These are basically noise modes, which exist because we have only 30 models for our PCA. (Numerical noise due to round-off error in the matrix operations is negligible.)

Measuring the short Q2–Q3 time delay, Fohlmeister et al. (2007) noted that substructure would be necessary for accounting for their observation. We have not considered separately the case where only this one time delay is known. We, however, see in figure 4 that only one of the

NTD eigenvalues is larger than the TD eigenvalues, indicating that only one of the measured time delays (surely the longer Q1–Q3 value) gives a substructure constraint. This does not mean that the Q2–Q3 measurement can be explained without substructure; only that the measurement on its own is not constraining what the substructure could be.

Figure 6 shows the variation in lens potential $t_{\rm grav}$ corresponding to the density variation from the main panel in figure 5. Here we can see the effect of measuring the delays between Q1, Q2, and Q3. The NTD models have $t_{\rm grav}$ between Q1 and Q2 varying by about 0.5 yr, that is, varying by 25% of the measured value. Between Q2 and Q3, the time delay varies by about 0.3 yr in the NTD models, which is more than twice the measured delay between this pair. Between Q3 and Q4 there is little variation. This does not mean that the NTD models have little variation in the Q3–Q4 delay; it just means that similar variation is present in the TD models, and hence not ruled out by the Q1–Q2–Q3 measurements. All these are what one would have expected from figure 3. What is surprising, however, is that the largest variation ruled out by the time delays is the blob 5″ west of the cluster centre, not near any of the images.

Figure 7 shows the enclosed mass and its variation. The enclosed mass is comparable with figure 3 of Oguri (2010), as was expected. As also evident from figure 5, the NTD mode is not like a global change of steepness—the steepness appears to have been already constrained by the image data, because of the redshift contrast between the Q, A, and B systems. So the time delays are giving us information not on the steepness but about the shape of the profile. The NTD models allow for strongly east/west ellipticity in the central region, changing to a more north/south elongation further out, but the time delays force the model rounder and less lopsided.

6 Discussion

This paper expresses the information that comes from lensing time delays in a way that is orthogonal to other lensing observables. We used the cluster SDSS J1004+4112 because of the richness of strong-lensing information. SDSS J1029+2623 (Fohlmeister et al. 2013; Oguri et al. 2013) would also be interesting for a similar study, as would MACS J1149+2223 (Kelly et al. 2014; Oguri 2014) if time delays for the recently discovered supernova can be measured. The main results are shown in figures 5 and 6. The interpretation is very preliminary, because a new time delay is expected soon, but nonetheless shows two interesting features.

First, it is remarkable how a small mass redistribution can produce a large difference in time delays. As shown in figure 5, the main mass-redistribution ruled out by the two published time delays is a small blob that is only $\sim 1\%$ of the total mass in the strong-lensing region, and yet this redistribution can change a time delay by 50%. It appears that time delays are providing substructure constraints at the 1% level.

Secondly, the mass redistribution is not mainly near the images, but mainly in another region of the cluster. Overall, the time delays reduce the allowed lopsidedness of the cluster, but it is intriguing that the main redistribution appears to shift mass from the neighbourhood of one cluster galaxy to the vicinity of another cluster galaxy.

The expected third time-delay measurement in this cluster will be very interesting. As seen in figure 1, the two time delays we used in our analysis are between Q1 and Q2, and between Q1 and Q3. All of these images are towards the south and southeast of the cluster. The new measurement would involve Q4, which lies to the north.

Thus we conclude that strong lensing with time delays provides important constraints on the distribution of matter near the centre of the lensing cluster, regions not accessible

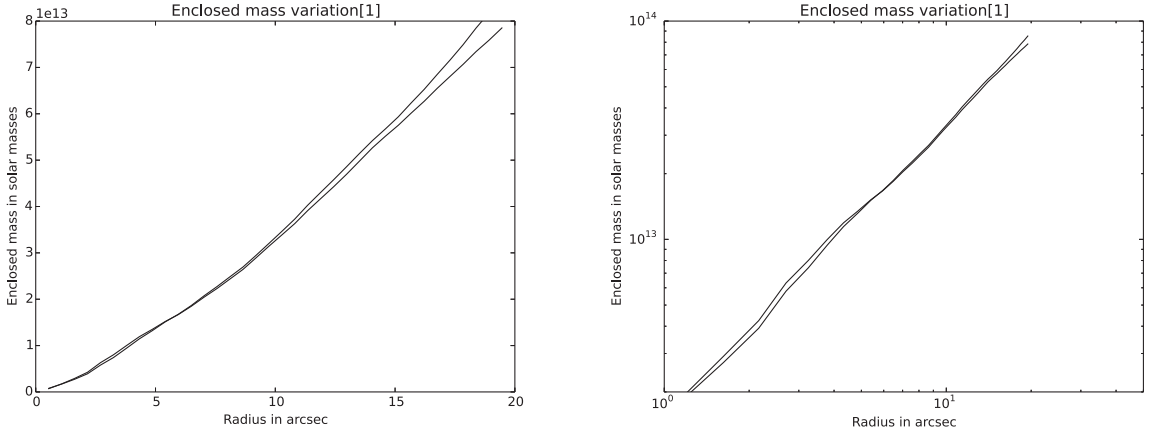

Fig. 7. Variation of enclosed mass of the reference model around the largest NTD mode (shown in figure 5) as a function of projected radius. Linear scales on the left, and log scales on the right.

to weak lensing or flexion. Additionally, time delays may provide new information on the distribution of matter in the densest regions of clusters and indirectly on the role of AGN feedback, adiabatic contraction, and other dynamical processes.

References

AbdelSalam, H. M., Saha, P., & Williams, L. L. R. 1998a, MNRAS, 294, 734

AbdelSalam, H. M., Saha, P., & Williams, L. L. R. 1998b, AJ, 116, 1541

Blandford, R., & Narayan, R. 1986, ApJ, 310, 568

Blumenthal, G. R., Faber, S. M., Flores, R., & Primack, J. R. 1986, ApJ, 301, 27

Coles, J. 2008, ApJ, 679, 17

Dehnen, W. 2005, MNRAS, 360, 892

Falco, E. E., Gorenstein, M. V., & Shapiro, I. I. 1985, ApJ, 289, L1

Fohlmeister, J., et al. 2007, ApJ, 662, 62

Fohlmeister, J., Kochanek, C. S., Falco, E. E., Morgan, C. W., & Wambsganss, J. 2008, ApJ, 676, 761

Fohlmeister, J., Kochanek, C. S., Falco, E. E., Wambsganss, J., Oguri, M., & Dai, X. 2013, ApJ, 764, 186

Gnedin, O. Y., Kravtsov, A. V., Klypin, A. A., & Nagai, D. 2004, ApJ, 616, 16

Halkola, A., Hildebrandt, H., Schrabback, T., Lombardi, M., Bradač, M., Erben, T., Schneider, P., & Wuttke, D. 2008, A&A, 481, 65

Hammer, F. 1991, ApJ, 383, 66

Inada, N., et al. 2003, Nature, 426, 810

Inada, N., et al. 2005, PASJ, 57, L7

Inada, N., Oguri, M., Falco, E. E., Broadhurst, T. J., Ofek, E. O., Kochanek, C. S., Sharon, K., & Smith, G. P. 2008, PASJ, 60, L27

Kawano, Y., & Oguri, M. 2006, PASJ, 58, 271

Kelly, P. L., et al. 2014, arXiv:1411.6009

Kneib, J.-P. 2002, in The Shapes of Galaxies and their Dark Halos, ed. P. Natarajan (Singapore: World Scientific), 50

Liesenborgs, J., & De Rijcke, S. 2012, MNRAS, 425, 1772

Liesenborgs, J., De Rijcke, S., & Dejonghe, H. 2006, MNRAS, 367, 1209

Liesenborgs, J., De Rijcke, S., Dejonghe, H., & Bekaert, P. 2007, MNRAS, 380, 1729

Liesenborgs, J., De Rijcke, S., Dejonghe, H., & Bekaert, P. 2008, MNRAS, 386, 307

Liesenborgs, J., De Rijcke, S., Dejonghe, H., & Bekaert, P. 2009, MNRAS, 397, 341

Lubini, M., Sereno, M., Coles, J., Jetzer, P., & Saha, P. 2014, MNRAS, 437, 2461

Martizzi, D., Teyssier, R., & Moore, B. 2013, MNRAS, 432, 1947

Mashchenko, S., Couchman, H. M. P., & Wadsley, J. 2006, Nature, 442, 539

Mohammed, I., Liesenborgs, J., Saha, P., & Williams, L. L. R. 2014a, MNRAS, 439, 2651

Mohammed, I., Martizzi, D., Teyssier, R., & Amara, A. 2014b, MNRAS submitted (arXiv:1410.6826)

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Oguri, M. 2007, ApJ, 660, 1

Oguri, M. 2010, PASJ, 62, 1017

Oguri, M. 2014, MNRAS in press (arXiv:1411.6443)

Oguri, M., et al. 2004, ApJ, 605, 78

Oguri, M., et al. 2013, MNRAS, 429, 482

Ota, N., et al. 2006, ApJ, 647, 215

Paraficz, D., & Hjorth, J. 2010, ApJ, 712, 1378

Rathna Kumar, S., Stalin, C. S., & Prabhu, T. P. 2014, A&A submitted (arXiv:1404.2920)

Read, J. I., & Gilmore, G. 2005, MNRAS, 356, 107

Refsdal, S. 1964, MNRAS, 128, 307

Refsdal, S. 1966, MNRAS, 132, 101

Saha, P., Coles, J., Macciò, A. V., & Williams, L. L. R. 2006, ApJ, 650, L17

Saha, P., & Read, J. I. 2009, ApJ, 690, 154

Saha, P., & Williams, L. L. R. 1997, MNRAS, 292, 148

Saha, P., & Williams, L. L. R. 2006, ApJ, 653, 936

Saha, P., Williams, L. L. R., & Ferreras, I. 2007, ApJ, 663, 29

Schaller, M., et al. 2014, MNRAS submitted (arXiv:1409.8617)

Schneider, P., & Sluse, D. 2013, A&A, 559, A37

Sereno, M., & Paraficz, D. 2014, MNRAS, 437, 600

Sharon, K., Gladders, M. D., Rigby, J. R., Wuyts, E., Bayliss, M. B., Johnson, T. L., Florian, M. K., & Dahle, H. 2014, ApJ, 795, 50

Sharon, K., et al. 2005, ApJ, 629, L73

Suyu, S. H., et al. 2014, ApJ, 788, L35

Teyssier, R., Pontzen, A., Dubois, Y., & Read, J. I. 2013, MNRAS, 429, 3068

Williams, L. L. R., & Saha, P. 2004, AJ, 128, 2631

Young, P., Gunn, J. E., Oke, J. B., Westphal, J. A., & Kristian, J. 1981, ApJ, 244, 736

Zhao, H., & Qin, B. 2003, ApJ, 582, 2

6

Paper5: Quantifying sub-structures in lensing clusters

Numerical simulations have provided an insight to the theory of structure formation. Large collapsed objects, like cluster of galaxies, form by a series of mergers of small mass systems, like galaxies or group of galaxies, and the associated gas, dark-matter halos etc. The abundance of sub-structures in a merging system should be quantified to study the merging stage of the cluster. This evaluation is trivial using simulations. The observed macroscopic properties of galaxies and galaxy clusters, for example gas mass, bulge sizes etc., can be well reproduced using the current generation hydrodynamical simulations. However, a more ambitious comparison would be the total mass distribution in the clusters. As the total mass distribution in an ongoing merger is the result of an initial random field evolved with deterministic forces, only statistical properties can be derived, and not the actual distribution. The derived statistical quantities like radial density profile, sub-structure mass function etc., can be compared to those of the observed clusters to validate the simulation ingredients.

Tracing the mass distribution in observed clusters is a non-trivial task, as the mass is not an observable, which light is. Gravitational lensing, particularly strong lensing, is an unbiased tool to estimate the total mass distribution. However, the resolution of the reconstructed mass maps depends on the priors and the lensing data. In order to have unbiased estimates, priors must be minimal, and the total mass distribution must be inferred using lensing data alone. Therefore, we need non-parametric mass reconstruction methods without light-traces-mass (LTM) assumption, and higher density of lensed images of the background sources, in order to correctly model the mass distribution of the lens (or the cluster).

Such large quantity of lensing data have recently been provided by the Hubble

Space Telescope (HST) under the Hubble Frontier Field (HFF) discretionary program. The HFF consists of six massive clusters of galaxies in redshift range 0.3 to 0.5, each containing 200 lensed, along with the good estimates of their redshifts.

In this paper, we produced mass models for the six HFF clusters using the pre-HFF data, and a mass model using HFF data for one of them. We report the gain in the spatial resolution of mass maps using HFF data over pre-HFF data. To make these mass maps, we use GRALE, a strong gravitational lensing non-parametric mass reconstruction technique, without assuming any light information from the clusters, except for the redshift. The mass distribution of all HFF clusters show elongation, multiple-cores, and many sub-structures, indicating a recent major merger. Therefore, extracting radial profiles of the clusters is not very encouraging. Also, because gravitational lensing can only estimate the sky-projected 2D mass distribution, many sub-structures are erased, and an estimate of the sub-structure mass function cannot be derived without LTM assumption.

We proposed the power spectrum of the 2D mass distribution as an estimator for the sub-structure. We measured this quantity for all HFF clusters, and found large power at small scales for clusters at low redshift or with larger number of lensed images. We made similar measurements of the power spectrum from simulated clusters, both dark-matter only and hydrodynamical, and found that the average power in simulated clusters is larger than that in the observed clusters at small scales. We discussed that the possible reasons for this could be: (i) limited lensing data, (ii) contrast in redshift between observed clusters and simulations, or (iii) lack of physics in simulations.

Role: For this project, I used the mass maps made by Professor Liliya. L. R. Williams and Kevin Sebesta. I proposed the idea to measure the 2D power spectrum of the mass distribution. I started by doing a null test, which shows equal power at all scales if there is only noise in the field. I, then, generated three random cluster fields for a given power spectrum using a code written by Dr. Prasenjit Saha. I measured the power spectrum of these simulated fields, and compared it to the original power spectrum; a good agreement was found. I accumulated simulated clusters data from Davide Martizzi, and measured the power spectra of all 66 clusters, each for DMO as well as hydrodynamical simulations. Finally, I made average power spectra of simulated and observed clusters, and drew a comparison.

This paper has been submitted to *Monthly Notices of the Royal Astronomical Society* (MNRAS), and is currently under review.
Arxiv: `http://arxiv.org/abs/1507.01532`

Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 7 July 2015 (MN LATEX style file v2.2)

Quantifying substructures in *Hubble Frontier Field* clusters: comparison with ΛCDM simulations

Irshad Mohammed⋆,[1,2] Prasenjit Saha[1,2], Liliya L. R. Williams[3], Jori Liesenborgs[4] and Kevin Sebesta[3]

[1] *Physik-Institut, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
[2] *Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
[3] *School of Physics & Astronomy, University of Minnesota, 116 Church Street SE, Minneapolis, MN 55455, USA*
[4] *Expertisecentrum voor Digitale Media, Universiteit Hasselt, Wetenschapspark 2, B-3590, Diepenbeek, Belgium*

7 July 2015

ABSTRACT

The Hubble Frontier Fields (HFF) are six clusters of galaxies, all showing indications of recent mergers, which have recently been been observed for lensed images. As such they are the natural laboratories to study the merging history of galaxy clusters. In this work, we explore the 2D power spectrum of the mass distribution $P_{\rm M}(k)$ as a measure of substructure. We compare $P_{\rm M}(k)$ of these clusters (obtained using strong gravitational lensing) to that of ΛCDM simulated clusters of similar mass. To compute lensing $P_{\rm M}(k)$, we produced free-form lensing mass reconstructions of HFF clusters, without any light traces mass (LTM) assumption. The inferred power at small scales tends to be larger if (i) the cluster is at lower redshift, and/or (ii) there are deeper observations and hence more lensed images. In contrast, lens reconstructions assuming LTM show higher power at small scales even with fewer lensed images; it appears the small scale power in the LTM reconstructions is dominated by light information, rather than the lensing data. The average lensing derived $P_{\rm M}(k)$ shows lower power at small scales as compared to that of simulated clusters at redshift zero, both dark-matter only and hydrodynamical. The possible reasons are: (i) the available strong lensing data are limited in their effective spatial resolution on the mass distribution, (ii) HFF have yet to build the small scale power they would have at $z \sim 0$, or (iii) simulations are somehow overestimating the small scale power.

Key words: gravitational lensing: strong, galaxies: clusters: individual: Abell 2744, Abell 370, Abell S1063, MACS J0416.1+2403, MACS J0717.5+3745, MACS J1149.5+2223

1 INTRODUCTION

Clusters of galaxies are the largest self-gravitating objects in the universe. Their ultimate origins must lie in some process of quantum fluctuations in an expanding universe (Schrodinger 1939; Harrison 1967; Fedderke et al. 2015), but the earliest observable precursors of galaxy clusters are fluctuations in the cosmic microwave background. CMB fluctuations on the scale of individual clusters would be at $l \sim 10^4$, far in the diffusion-damping tail (Silk 1968) and so far barely accessible observationally (Reichardt et al. 2009), but their subsequent growth through linear gravitational instability is straightforward.

The early observable proto-clusters (e.g, Toshikawa et al. 2014) are, however, far beyond the baby fluctuations of the linear regime of gravitational instability — in the well-known toy model of spherical collapse, the dynamical outer boundary of a cluster is at an overdensity of $\simeq 200$. In this regime some analytical methods based on generalising spherical collapse are available (Press & Schechter 1974; Bond et al. 1991; Seljak 2000) but for the most part theoretical study depends on numerical simulations.

With the recent developments in computational resources, it is now possible to simulate dark matter and gas in cosmological volumes with good resolution. In such simulations it is possible to track the evolution and mergers of small systems to form large collapsed clusters of galaxies. The agreement between simulations and observations have been improving for various macroscopic proper-

⋆ irshad@physik.uzh.ch

ties of galaxies, such as intergalactic gas, bulge sizes etc., (Klypin et al. 2011; Angulo et al. 2012; Alimi et al. 2012; Skillman et al. 2014; Agertz et al. 2011; Vogelsberger et al. 2013, 2014; Marinacci et al. 2014; Schaye et al. 2015). To study individual objects in more detail, hydrodynamical zoom-in simulations can be used (Feldmann & Mayer 2015; Fiacconi et al. 2015).

A more ambitious comparison between simulations and observations would be that in terms of the distribution of mass in the clusters. This is non-trivial for two reasons:

- First, mass is not an observable; what we observe on the sky is light, mass can only be inferred. Tracing the mass typically involves additional assumptions, such as that galaxies sit in the potential wells of the dark-matter, light traces mass with some scaling parameters, etc.
- Second, since individual halos, galaxies or clusters of galaxies are the outcome of gravitational collapse and various baryonic processes of a random field (initial density field), it is not possible to compare their mass distribution or clustering properties directly. The best one can hope for is to compare them statistically.

The first problem could be solved with the help of gravitational lensing, which is sensitive only to mass. Well-known examples of light not tracing mass, revealed by lensing, are the clusters ACO 2744 (Merten et al. 2011) and recently ACO 3827 from Massey et al. (2015) (see also Williams & Saha 2011; Mohammed et al. 2014). But simulations of cluster formation cannot be fitted so as to reproduce detailed properties of individual clusters.

The second problem could be handled by identifying robust statistical properties of the clusters. The simplest quantity is the radial density profiles. Newman et al. (2013b,a) studied the average density profiles of the lensing clusters and compared them to the simulated clusters. However, this quantity can only be measured if the cluster is virialised, definitely not for an ongoing merger. High-mass clusters in simulations do not usually virialise until $z \sim 0.3$; before that they often show elongation, multiple cores and many substructures indicating a recent merger, as do observed lensing clusters.

Another very popular statistic is the mass function; counting the number of sub-halos as a function of their masses (Natarajan & Springel 2004; Natarajan et al. 2007; Atek et al. 2015). However, measuring or even identifying substructures in lensing clusters has so far been possible only under the assumption that light traces mass. Without this assumption, the mass function does not seem a viable statistic with lensing, especially since lensing gives information about the sky-projected mass and projection may wash out substructures.

In this work, we propose a different strategy, which is to use a two-dimensional power spectrum as a basis for comparing lensing clusters with simulations. In section 2, we define such a power spectrum, which is normalised differently from the usual cosmological power spectrum, and has dimensions of mass-squared. In section 3, we briefly describe GRALE, a free-form lens-reconstruction technique and code. In section 4 we apply GRALE to reconstruct the six clusters of the Hubble Frontier Field from strong-lensing data, and calculate their power spectra. Then in section 5 we compare the clusters with each other and with ΛCDM simulations, both dark-matter only and hydrodynamical. Finally, section 6 has general discussion and suggestions for future work.

2 SUBSTRUCTURE POWER SPECTRUM

Because it is impossible to directly compare mass maps of observed and simulated clusters comparison of the corresponding properties like number of galaxies, 1D density profiles, concentration, mass function of the structures/substructures, temperature profiles etc. are more useful. However, none of these quantify the clustering properties of the halo/cluster which contain important information about its merging history and evolution. So it is interesting to study statistically the cluster mass distribution. The simplest statistic is a two-point function. Cosmologically, a two-point function gives an excess probability that a local density peak exists close to another peak, as a function of the separation between the two. It can be well studied in the Fourier space as the power spectrum. Generally, a power spectrum $P(k)$ of a 3 dimensional matter density field is defined as:

$$P(k) = \langle |\tilde{\delta}(k)|^2 \rangle \tag{1}$$

where, $\tilde{\delta}(k)$ is the Fourier transform of the over-density at position corresponding to the wave-vector k.

In this work we measure the 2D power spectrum from the projected mass distribution of the halos:

$$P_{\rm M}(k) = \langle |\tilde{M}_{2D}(k)|^2 \rangle \tag{2}$$

where, $\tilde{M}_{2D}(k)$ is the Fourier transform of the 2D projected mass element at position corresponding to the wave-vector k.

$$\tilde{M}(\vec{k}) = \int \Sigma(\vec{x}) e^{i\vec{k}\cdot\vec{x}} d^2\vec{x} \tag{3}$$

This form also gives a natural normalisation of the function to be the square of the total mass of the halo.

This is not the only way to describe the statistical properties of a density peak, For example, Hezaveh et al. (2014), using the halo model approach, define

$$P_{\rm 2D}(k) = \int \frac{dn}{dM} |\tilde{\kappa}(k)|^2 dM, \tag{4}$$

where $P_{\rm 2D}(k)$ is the so-called one-halo term in two dimensions, $\frac{dn}{dM}$ is the differential mass function (number density of halos per unit mass) and $\tilde{\kappa}(k)$ is the Fourier transform of the convergence (the normalised surface mass density). The framework is based on the assumptions of virialised halos, and a functional form for its mass function and radial density profile of each halo. The total contribution to the power spectrum comes by adding the correlation between different halos (the two-halo term) and the correlation between matter within the same halo (the one-halo term). Within a halo, there are also many substructures. If one wants to study the clustering properties of matter within a gravitationally bound system like a galaxy cluster (or a merging system of galaxies), the two-halo term can be neglected as all the structures are moving within the gravitational potential of the host and are indifferent to each other's gravity. However, one-halo term still exists, and at large scales it contributes as the Poisson distribution of the substructures which drops at the size of the largest substructure. However,

this can only apply if the cluster (or the halo) is virialised so that a smooth density profile can be subtracted in order to see the correlation between the residual field. Now, this is not a good assumption overall, especially during merging. As most of the massive galaxy cluster lenses at redshift $> 0.3-0.4$ are not virialised systems, this form of the power spectrum is not intuitive. Therefore, for a general system, including ongoing mergers, recent mergers and virialised halos, the correlation between the distribution of matter inside the halos must be studied without such assumptions.

With our definition, if we have a field enclosing only one halo with its substructures, and we compute $P_{\rm M}(k)$ at scales larger than the size of the halo, for all those scales one gets the same mass and hence the power above all those scales will be a constant, mimics the Poisson noise of the substructures in the halo model picture. Further this function drops at scales smaller than the size of the halo, which is the largest structure in the field, consistent with the halo model intuition. For a virialised halo, this function will be smooth and a power law for all scales smaller than the size of the halo. However, if the halo is not spherically symmetric and in a merging stage the fluctuations in the function resemble variation in its power at various scales. If one identifies a large ensemble of halos (let's say in simulations) of nearly the same mass and the same epoch and average all the power spectra, one will get an unbiased trend of the clustering properties of such halos statistically. However, studying individual systems is like studying a random realisation of a halo which underlies a mean power spectrum and the recovered power spectrum from such a halo should be statistically consistent with the mean.

To ensure this feature, we performed a test. We generated three density fields which are the random realisation of the same power spectrum ($\propto k^{-3}$ at small scales) as shown in figure 1. The three density grids enclosed structures and some substructures and look very different morphologically, but as they are generated from the statistics, they should show the same power spectrum. In bottom-right of figure 1 we show the ratio of the recovered power spectrum from each of the grids to the input power spectrum. The slope of the recovered power spectrum is calculated by fitting the power law towards the larger values of k, between $6.8\times10^{-4}-4\times10^{-3}{\rm pc}^{-1}$. All power spectra give nearly the same slope as the input one with an error of about 10-20%. This test shows that given a density field, the correct mean statistic can be obtained within a reasonably small error.

For a pure noise field we expect same power at all scales. We performed this test as well (but we don't show the result in a figure form), and verified that the power spectrum is flat.

The fluctuations in mass distribution at various scales can be directly inferred as the presence/absence of substructures. Larger power at small scale indicates the presence of local density peaks, however, a flat power at all scales indicates similar structures at all scales or just the noise. Also smoother distribution of the matter gives smaller power whereas sharp peaks correspond to larger power. Therefore, we expect the power at small scales to increase as the merger progresses, and reach the highest power when the system is virialised. In other words, one should expect the power on small scales to decrease with redshift. This form of the power spectrum can be used in order to compare halos statistically, within simulations or observations or across the two.

3 LENS RECONSTRUCTION METHOD

In GRALE (Liesenborgs et al. 2007, 2008) the mass distribution is free-form and consists of a superposition of a large number of adjustable components (several hundred Plummer lenses). The distribution of these Plummer lenses is adaptively determined by GRALE using a multi-objective genetic algorithm, to optimise two fitness measures.

(i) The *overlap fitness* quantifies the fractional overlap of the projected images of the same source back on to the source plane. If all images of a source back-project to exactly the same area on the source plane, the source fitness for that image system is perfect. More generally, the larger overlap the better the fitness. It is important to use the fractional overlap, as otherwise the fitness measure would be biased towards extreme magnifications.

(ii) The *null fitness* is a penalty for any spurious images implied by a model where none are present in the observations. The penalty is applied only in regions where the data make it clear that no images are present; extra images are allowed in regions that are difficult to observe because, for example, of the presence of nearby bright galaxies. Among lens-modelling techniques, GRALE is unique in being able to exploit the *absence* of images as useful information.

Further fitness measures can be defined and incorporated into GRALE, in particular, time delays (Liesenborgs et al. 2009; Mohammed et al. 2015). No information about the light from the cluster is used — mass-traces-light assumptions are completely absent.

While the genetic algorithm optimizes the properties and placement of the component Plummer lenses, the user still needs to specify a range for the allowed number of components. This is, in effect, an overall resolution of the mass map. To set this effective resolution, we run the mass-reconstruction process first at coarse resolution, then finer, and then coarse again, and choose an optimal trade-off between fitness and resolution — see Mohammed et al. (2014) for details and tests of this strategy. Finally, the entire cycle is repeated 30 times to generate an ensemble of mass reconstructions.

4 HUBBLE FRONTIER FIELDS

Hubble Frontier Fields (HFF) survey (PI: J.Lotz, HST 13498) is a three year Director's discretionary time program that devotes a total of 840 orbits to six galaxy clusters plus the accompanying parallel fields. Each field is observed in three HST optical bands (ACS F435W, F606W and F814W) and 4 infra-red bands (WFC3-IR F105W, F125W, F140W, and F160W). It is the single most ambitious commitment of HST resources to the exploration of the distant universe through the power of gravitational lensing by massive galaxy clusters. All six clusters are early or intermediate stage mergers at $z = 0.3$–0.55, with significant elongation and hence non-trivial mass distribution. Each had about 10–20 multiply imaged background sources discovered with pre-HFF

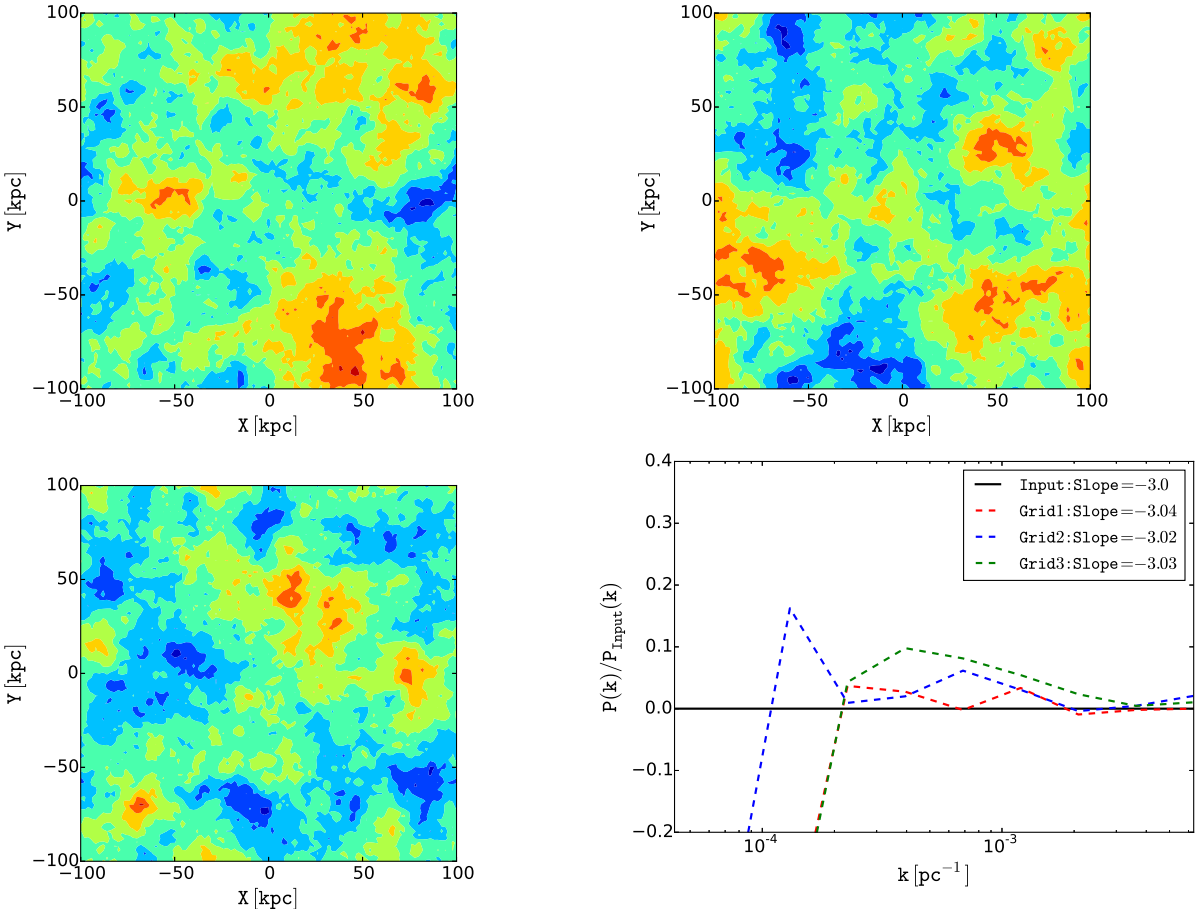

Figure 1. The first three plots (in reading order: Grid1, Grid2, Grid3) are the three density map realizations generated from the same power spectrum with k^{-3} towards large k. The bottom-right panel shows the comparison of the recovered power spectrum from each density grid to the input power.

HST data. Five independent teams were tasked with making mass and magnification maps for these six clusters. In the analysis of the present paper we use mass maps presented by our team, made with pre-HFF lensing data on six clusters as well as one map made with post-HFF, for cluster MACS 0416. Table 1 shows the necessary information about the lensing data in each of the HFF clusters.

4.1 MACS J0416.1-2403

MACS J0416.1-2403 (MACS0416 henceforth) is a strong gravitational lens at redshift nearly 0.396. It is an elongated galaxy cluster, most probably a recent merger or a pre-merger (Ogrean et al. 2015), and hence a non-trivial mass distribution. It was first identified by MAssive Cluster Survey (MACS; Ebeling et al. 2010). Based on its double-peaked surface brightness in the X-rays (Mann & Ebeling 2012), its recent merger stage is confirmed. It is amongst the five high magnification clusters in the CLASH (Cluster Lensing and Supernovae survey with Hubble) project (Postman et al. 2012). Detailed mass maps were provided by Zitrin et al. (2013) using CLASH data, and Jauzac et al. (2015) using HFF data.

We made two reconstructions of this lens using GRALE, one with pre-HFF data (total 40 images from 13 sources), using data from Merten et al. (2011), as well as data provided by J. Richard and D. Coe, and one with HFF data (total 149 images from 57 sources) which were identified by Jauzac et al. (2015). The two reconstructed mass models—with each being an average of 30 realisations—are shown in figure 2 (left column) with respective power spectrum (highlighted in the right column). Both mass models are very similar at large scales, which is also evident from the power spectrum, and only differ in the details at small scales. The two mass clumps and the elongation can be identified in both models. However, the mass models with HFF data (bottom row), shows larger power or much detailed substructures at small scales. This is expected as HFF data has many more lensed images than pre-HFF data and hence provides additional lensing constraints at small scales, which leads to increased power at larger k's.

4.2 Abell 2744

Abell 2744 is a massive galaxy cluster at redshift 0.308 and an active merger. It has been studied in various

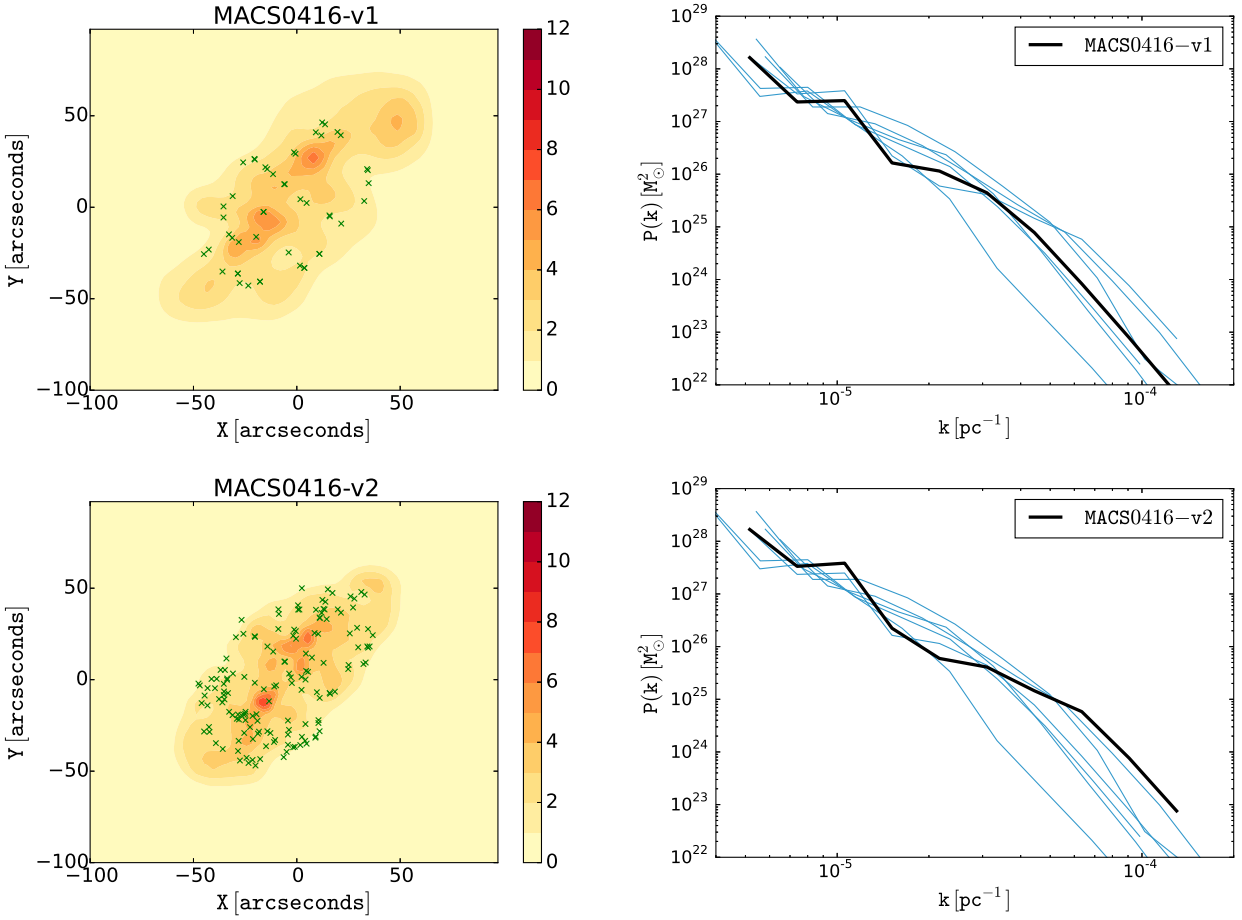

Figure 2. Left: the density map for MACS0416 cluster using pre-HFF data (top-row) and HFF data(bottom-row). The crosses mark the positions of the lensed images. Right: The highlighted (black) corresponding power spectrum of the cluster on the left with all other HFF clusters' power spectra in the background.

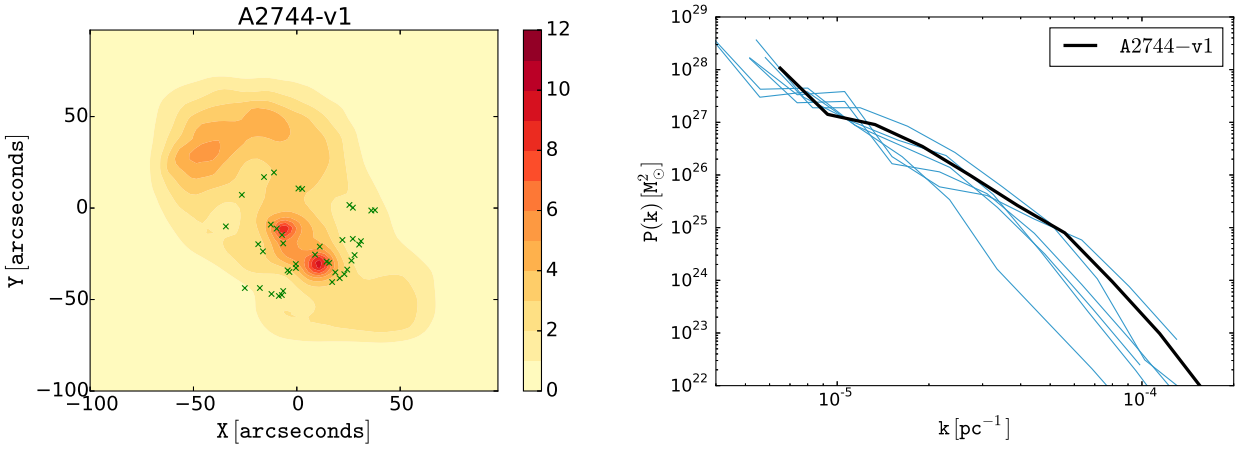

Figure 3. Left: the density map for Abell 2744 cluster using pre-HFF data. The upper left clump could well be an artefact as there are no images there. Right: The highlighted (black) corresponding power spectrum of the cluster on the left with all other HFF clusters' power spectra in the background.

Cluster	Data	Redshift	Number of images	Number of sources	Redshift range of sources
MACS-J0416	pre-HFF	0.396	40	13	1.82-3.25
MACS-J0416	post-HFF	0.396	149	57	1.00-5.90
Abell 2744	pre-HFF	0.308	41	12	2.00-4.00
Abell 370	pre-HFF	0.375	36	11	0.73-3.00
AS-1063	pre-HFF	0.348	37	13	1.22-6.09
MACS-1149	pre-HFF	0.543	32	11	1.23-3.80
MACS-0717	pre-HFF	0.545	23	7	1.80-2.91

Table 1. Lensing data from different HFF clusters.

wavelengths, for example, it has an extended radio halo (Giovannini et al. 1999; Govoni et al. 2001), X-ray emission (Allen 1998; Ebeling et al. 2010; Kempner & David 2004; Owers et al. 2011) and various substructures have been identified in the optical observations (Girardi & Mezzetti 2001; Boschin et al. 2006). Shan et al. (2010) shows a significant offset between the dark-matter component (from lensing) and baryonic matter (from X-rays observations). There is also a magnified singly imaged supernova in the background, at redshift 1.35 (Rodney et al. 2015). Various lensing analyses have been done (Smail et al. 1997; Allen 1998; Cypriano et al. 2004; Merten et al. 2011; Lam et al. 2014; Jauzac et al. 2014).

We present the mass model of Abell 2744 using GRALE with pre-HFF data (Merten et al. 2011; Richard et al. 2014) as well as data provided by J. Richard and D. Coe, for a total of 41 images from 12 background sources spread over a redshift range of 2 to 4. Figure 3 shows the reconstructed mass map and the corresponding power spectrum. It shows two very distinct blobs and the overall elongated structure, a morphology very likely for a major merger which is in fact confirmed by previous studies.

Due to very sharp mass peaks, this cluster, like MACS0416, also shows larger power at small scales.

4.3 Abell 370

Abell 370 is a strong gravitational lens at redshift 0.375 and hosts a giant gravitational arc at redshift 0.725. Due to its large Einstein radius, it is an ideal target to look for high redshift galaxies through its magnification, especially in high density regions. There are many published mass models (for example Soucail et al. 1987; Abdelsalam et al. 1998; Richard et al. 2010). We used pre-HFF data from the latter work, as well as data provided by D. Coe and A. Koekemoer, to reconstruct its mass distribution. Figure 4 shows the resulting mass map. The mass distribution is shallow but shows many substructures. The resulting power spectrum shows less power on small scales as compared to other HFF clusters which is reflected in shallower peaks in its mass distribution.

4.4 Abell S1063

Abell S1063 is a galaxy cluster at redshift 0.348. We used pre-HFF data identified by Richard et al. (2014); our reconstruction used a total of 37 images from 13 background sources.

Figure 5 shows the reconstructed mass map along with the corresponding power spectrum. It shows two mass clumps, which lead to larger power at intermediate scales, but due to the shallowness of the peaks, it loses power at small scales.

4.5 MACS J1149.5+2223

MACS J1149+2223 (MACS1149 hereafter) is an X-ray bright cluster at redshift 0.543. It has been studied by various authors for its rich strong lensing (Smith et al. 2009; Zitrin & Broadhurst 2009; Zitrin et al. 2011, 2015; Rau et al. 2014; Johnson et al. 2014; Sharon & Johnson 2015). There is also a multiply imaged supernova (Kelly et al. 2014, 2015) hosted by a face-on spiral galaxy at redshift 1.49.

Figure 6 shows the reconstructed mass model of MACS1149 using GRALE with pre-HFF data (total 32 images from 11 background sources) from Smith et al. (2009); Zitrin & Broadhurst (2009), and data provided by M. Limousin. The mass model favours two dominant peaks and one sub-peak, elongation and many substructures. The peaks are shallower than those in Abell 2744, also expected as it has nearly the same mass as Abell 2744 but higher redshift. The corresponding power spectrum is shown in the right column of the same figure.

4.6 MACS J0717.5+3745

MACS J0717.5+3745 (MACS0717 hereafter) is a strong gravitational lens at redshift 0.55, classified as the most dramatic merger in X-ray/optical analysis by Mann & Ebeling (2012).

We used pre-HFF data from Limousin et al. (2012); Richard et al. (2014) and Medezinski et al. (2013) to reconstruct its mass distribution using GRALE. Figure 7 shows the resulting mass map and the corresponding power spectrum. Due to the clear mass peaks and substructures, the power shows increased power at intermediate scales, however, due to very shallow peaks the power spectrum at small scales drop and is amongst the lowest of the HFF clusters.

5 COMPARING THE CLUSTERS

5.1 Comparing HFF clusters

We can now compare the clustering properties of HFF clusters using the lensing power spectrum $P_{\mathrm{M}}(k)$. The reconstructed mass distributions of the HFF clusters are morphologically different but show similar statistical structures

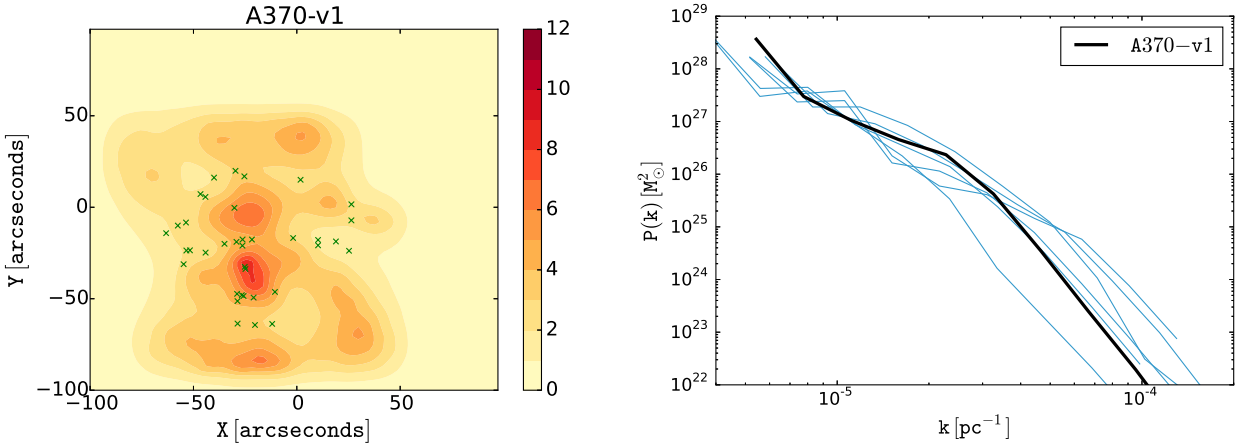

Figure 4. Left: the density map for Abell 370 cluster using pre-HFF data. Right: The highlighted (black) corresponding power spectrum of the cluster on the left with all other HFF clusters' power spectra in the background.

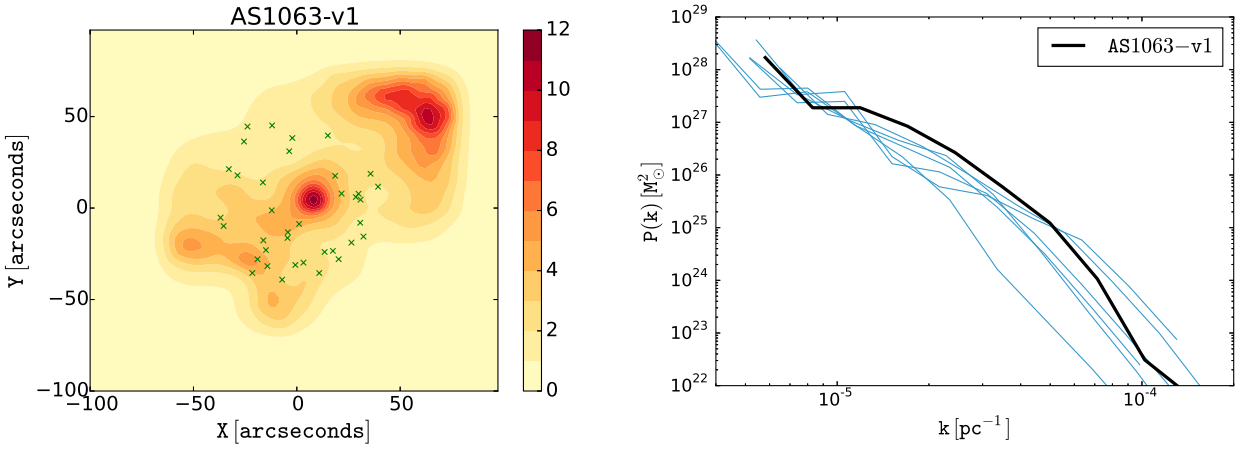

Figure 5. Left: the density map for AS1063 cluster using pre-HFF data. Right: The highlighted (black) corresponding power spectrum of the cluster on the left with all other HFF cluster's power spectrum in the background. Note that there are no images in the upper right corner of the map, so the massive clump in that region could well be an artefact.

which is very much evident by comparing the respective power spectra. Figure 8 shows all the power spectra for HFF clusters in thick lines along with their sample variance, which represents scatter across all available modes for a given k.

Let us first consider two reconstructions for MACS0416, using pre-HFF and HFF data. Both mass maps are very similar at large scales but differ in detail at small scales. The same effect can be noticed in the power spectrum (see figure 8), up to $k \sim 3 \times 10^{-5}\mathrm{pc}^{-1}$, both look nearly the same, but at smaller scales the mass map reconstructed with pre-HFF data starts to lose power. This is the consequence of the fact that HFF data has three times as many lensed images as pre-HFF data (see table 1), allowing for higher resolution in the reconstructed mass maps, which leads to higher power at small scales, where the difference in the two power spectra is almost an order of magnitude.

Power spectra for Abell 370 and MACS1149 look very close to each other. If we look at the corresponding mass maps, both show two dominant peaks with similar width, shallow as compared to the surroundings with similar smoothing. Similar arguments can be made for Abell 2744 and MACS0416-II clusters. The clusters show similar power at small scales which is expected from the very steep mass peaks in the two clusters.

It is important to note that the smallest scale (largest k) to be trusted depends on the density of images. For example, MACS0416-v2 has 6.5 times more images than MACS0717, (assuming roughly the same area for both in terms of arcsec^2), so the typical linear image spacing is 2.5 times larger in MACS0717, leading to poorer mass resolution. And, in fact, looking at figure 8, the k value where $P_\mathrm{M}(k)$ of MACS0717 begins to drop ($k \sim 2 \times 10^{-5}\mathrm{pc}^{-1}$) is about a factor of 2.5–3 smaller compared to that of MACS0416 ($k \sim 6 \times 10^{-5}\mathrm{pc}^{-1}$).

There is also a tendency for low redshift clusters to have higher power at small scales. Thus, Abell 2744 reconstructed with 41 images has larger power than MACS0416-v1 (40 im-

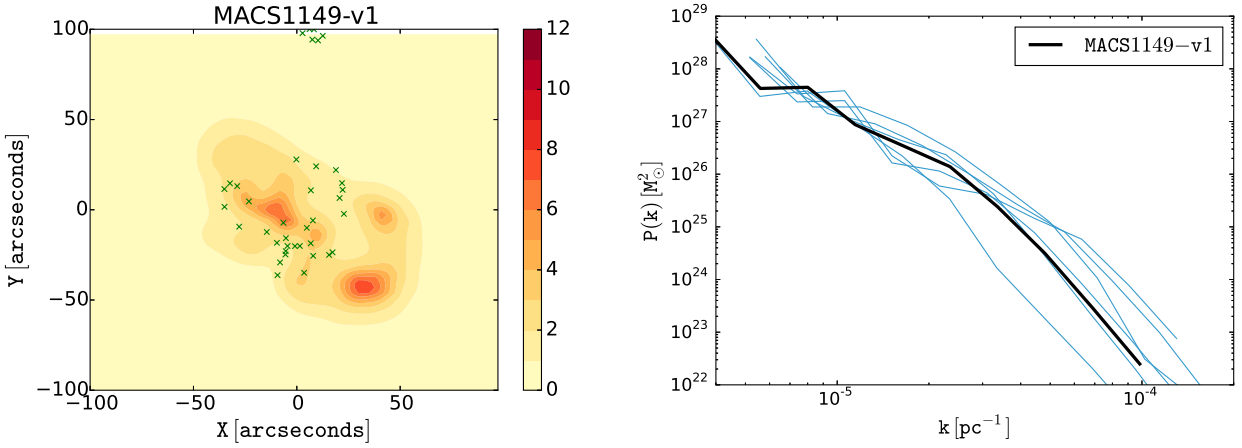

Figure 6. Left: the density map for MACS1149 cluster using pre-HFF data. Right: The highlighted (black) corresponding power spectrum of the cluster on the left with all other HFF clusters' power spectra in the background. There are no images beyond X=25", so the two clumps on the right could well be artefacts.

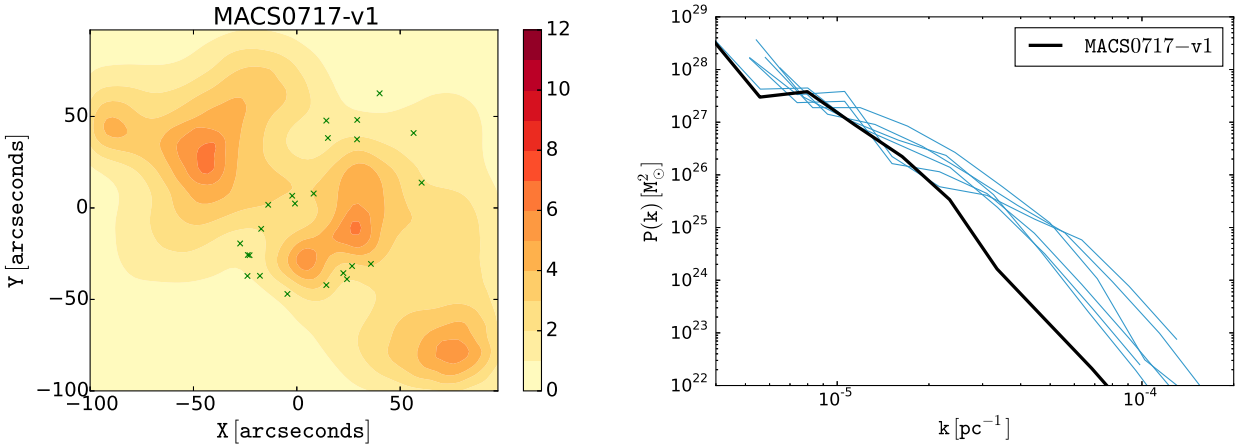

Figure 7. Left: the density map for MACS0717 cluster using pre-HFF data. The upper left and bottom right clumps could well be an artefact as there are no images there. Right: The highlighted (black) corresponding power spectrum of the cluster on the left with all other HFF cluster's power spectrum in the background.

ages) at small scales. Similarly, MACS1149 has less power at small scales than A370 and AS1063, again with similar numbers of lensed imaages. This is also intuitive as low redshift clusters have had more time to build small scale substructures. Therefore, the power at small scales maybe indicative of the age of the cluster.

5.2 ΛCDM simulation clusters

We used 22 clusters of galaxies from dark-matter only (DMO) simulations and further 22 from hydrodynamical simulations which include AGN feedback, in the mass range 1–3×$10^{14}\,M_\odot$ from Martizzi et al. (2014). The mass range is chosen to be comparable to that of the HFF clusters but the simulated clusters are all at redshift zero. We then projected them along the three axis, so a total 66 projected clusters in each case. In contract to HFF clusters, the simulation clusters are virialised but a few show more than one core. The measured power spectra are shown as thin lines in figure 8.

We also measured the mean power spectrum in each case, DMO and AGN, which are shown in black in left panel figure 9. The error bars show the standard deviation of $P_{\rm M}(k)$, including contributions from scatter in different modes at a given k as well as the scatter between different clusters. In the right panel of figure 9, we show the relative difference between the mean power in DMO and AGN simulated clusters. Comparing the mean power spectra of DMO and AGN, the latter shows a deficit in power at large scales and a boost at smallest scales. The deficit is nearly 10%. The AGN clusters have the influence of an AGN at the centre that drives the gas outside the cluster with its feedback process. Losing this mass results in the deficit of power at those scales. However, these AGN clusters contain a central stellar component in the form of the brightest cluster galaxy (BCG), which increases the mass at the very centre

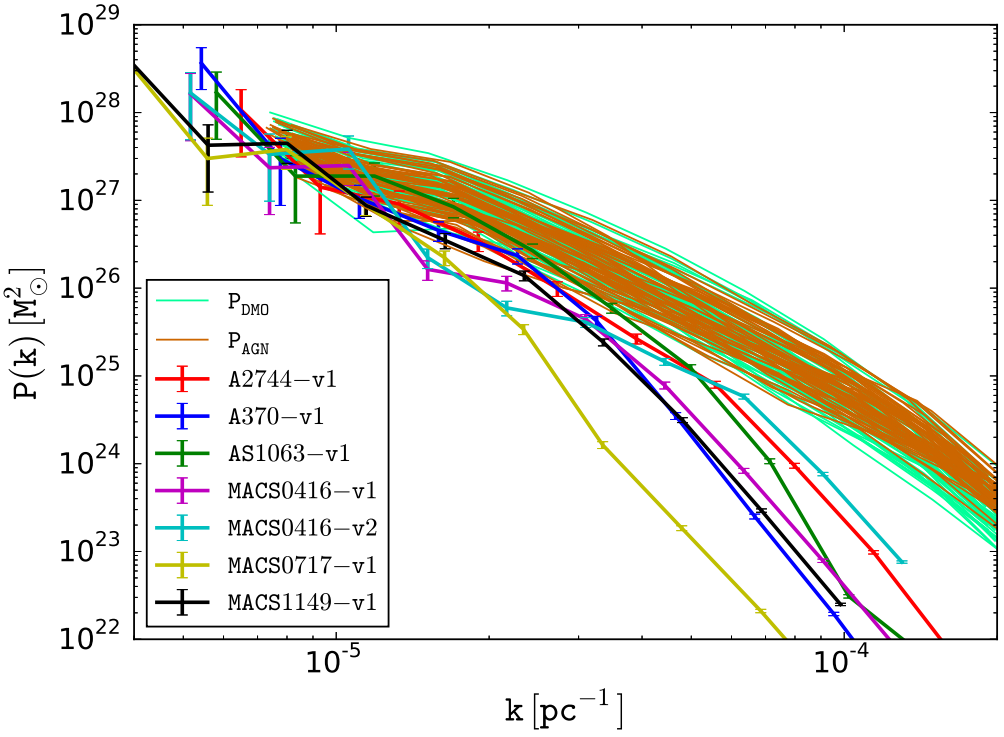

Figure 8. All power spectra: clusters in thick solid lines and simulations in thin solid lines.

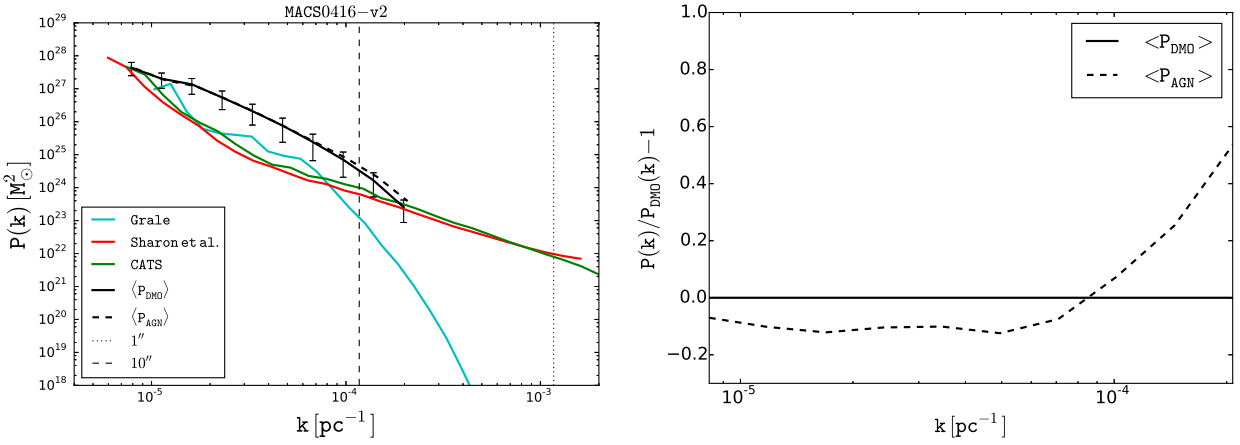

Figure 9. Left: Comparing the mean power spectrum for dark-matter only simulations (DMO), baryonic simulation (AGN) and different reconstructions using GRALE and LENSTOOL for MACS0416-v2 (using HFF data only). The vertical lines shows the 10 and 1 arc second scale at the cluster's redshift $z = 0.396$. Right: The relative difference between mean power spectra of DMO and AGN simulations.

of the cluster resulting in the boost in the power at smallest scales. The trend is systematically consistent with previous findings.

5.3 Comparison between HFF clusters and ΛCDM simulation clusters

In figure 8, we show $P_{\rm M}(k)$ for all six HFF clusters as well as for all the simulated clusters, both DMO and AGN. The lensing clusters systematically show less power at small scales compared to the simulated ones. We discuss the possible reasons for this in section 6.

5.4 Comparison with LTM models

In the left panel of figure 9, we compare the power spectrum measured from the mass maps reconstructed using GRALE to that made using other methods: Sharon et al and the

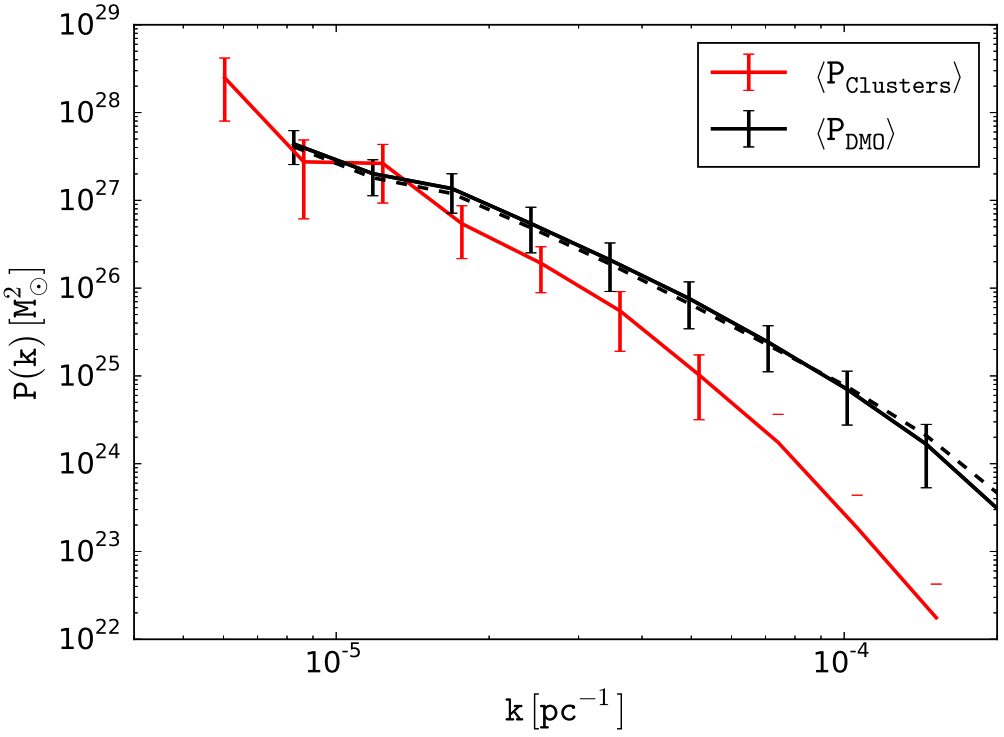

Figure 10. Comparing average power spectrum of HFF clusters to that of the simulated clusters.

CATS group.[1] Both use the LTM method LENSTOOL of Jullo et al. (2007) as the reconstruction technique. Both of these reconstructions show much larger power at small scales compared to the GRALE reconstruction. This is the result of the fact that LENSTOOL uses information from the individual cluster galaxies of the lens, and hence has much steeper density gradients at small scales which results in larger power. In contrast, free-form methods like GRALE base the mass distribution on the lensed images alone, without relying on visible galaxies. On the other hand, the power spectrum of parametrically reconstructed clusters is a mixture of priors and lensing information, and when the number of images is small, priors are the dominant contribution.

6 CONCLUSIONS AND DISCUSSIONS

In this article, we present free-form mass models of six Hubble Frontier Field clusters using strong gravitational lensing pre-HFF data. We also studied one of the six clusters, MACS0416, using both pre-HFF and HFF data (with three times more lensed images). The mass maps are made using a non-parametric lens inversion technique called GRALE. All the mass maps show elongation, multiple cores and many substructures in the mass distribution implying a recent major merger. The lensing data from recent HST-HFF observations is rich in lensing images allowing a very precise identification of substructures to be done with greater confidence. These mass maps can be used to study various aspects of the structure formation and merging stages of large collapsed clusters of galaxies.

We measured the power spectra of these sky-projected mass maps. Power spectra give a statistical description of the clustering properties of the mass distribution and encode information about the abundance of substructures and their contrast with the background. There are tentative indications that low redshift clusters have systematically higher power at small scales as compared to high redshift clusters, presumably because low redshift clusters are closer to their virial equilibrium than high redshift clusters. We argue—and illustrate using pre-HFF and HFF maps of MACS0416—that using a larger number of lensed images in a cluster leads to a better constrained mass distribution, and that the power spectrum can be recovered up to much smaller scales putting stronger constraints on our understanding of the substructures. On the other hand, parametric methods that explicitly include cluster galaxies in the models have much larger power on small scales even with fewer lensed images indicating that the mass maps are dominated by the priors, especially at small scales.

We compared the power spectrum of the HFF clusters with those from ΛCDM dark-matter only and hydrodynamical simulations at redshift zero. Figure 10 summarizes our comparison between HFF clusters and ΛCDM simulation clusters. We average the power spectra of the six HFF clusters and compare them with the average power spectra of the simulated clusters. The average power of the clusters is systematically steeper than that of the simulated clusters

[1] `https://archive.stsci.edu/prepds/frontier/lensmodels/`

and hence have less power at small scales. This may be due to one or more of the following reasons:

• The redshift of the HFF clusters are higher than the simulated clusters. We are comparing non-virialised HFF clusters with the virialised halos of the simulations. At later times, the mass distribution becomes more and more clumpy, and the substructures pull more mass from the background, which results in higher power at small scales. Therefore, it is possible that the lensing clusters in the local Universe may show similar power as the simulated clusters at redshift zero. Conversely, this can be checked by comparing the power spectrum of HFF clusters with that of simulated clusters of similar masses in the redshift range 0.3–0.5.

• A second possible reason relates to the data and method we are using to reconstruct the clusters. A larger number of lensed images gives additional constraints on the mass distribution, and this increased resolution leads to a boost in the power spectrum. This can be seen in figure 8 where we show the power spectrum of MACS0416 using both pre-HFF and HFF data: HFF data shows larger power at small scales as compared to pre-HFF data.

Figure 9 shows that the power at small scales also depends on the assumption of the reconstruction method. It is possible that GRALE does not have enough resolution and hence loses power at smallest scales. On the other hand LENSTOOL maps seem to have much more power than expected at those scales and are possibly dominated by priors.

• Finally there is the possibility that the simulations are lacking some physics that need to be taken into account in order to simulate realistic halos.

The second item above can be tested through a more extended pipeline, in which clusters are generated from state-of-the-art simulations, lensing data are generated from them, and then analysed by independent groups using different techniques. Such tests are currently in progress. It will be interesting to measure the power spectrum of the original mass distribution and the reconstructed one using different lens inversion methods. The current analysis expects that the power spectrum based on GRALE maps will match the true power spectrum at large scales and ultimately lose power depending upon the density of lensed images, whereas, LTM-based methods may continue having larger power even when the true power drops. We leave this analysis for future work.

Independently of lensing, it will also be interesting to study the evolution of the power spectrum as a probe of substructures and the merging history of large collapsed objects in simulations. For example, a cosmological simulation can be set up with a volume large enough to produce 20–50 clusters of galaxies, in a narrow mass range. The power spectra can be calculated for all the clusters in the mass range at different redshifts, and the evolution of the average power spectrum can be studied. This will give us an insight about the evolution of clustering properties in a merging system. We leave this analysis to future work.

7 ACKNOWLEDGEMENT

We would like to thank Davide Martizzi to provide the projected mass maps of the simulated clusters. IM would also like to thank Aurel Schneider for useful discussions about the topic.

LLRW is grateful to the Minnesota Supercomputing Institute for their computational resources and support.

We are grateful to a number of HFF Map Makers, Dan Coe, Johan Richard and Anton Koekemoer who made their image identifications and redshifts available to us, in some cases prior to publications. Some redshifts used were spectroscopic, some photometric and some were predicted by the lens models. The relevant papers from which image information was taken are cited in the text.

REFERENCES

Abdelsalam H. M., Saha P., Williams L. L. R., 1998, MNRAS, 294, 734
Agertz O., Teyssier R., Moore B., 2011, MNRAS, 410, 1391
Alimi J.-M. et al., 2012, ArXiv e-prints: 1206.2838
Allen S. W., 1998, MNRAS, 296, 392
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Atek H. et al., 2015, ApJ, 800, 18
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Boschin W., Girardi M., Spolaor M., Barrena R., 2006, A&A, 449, 461
Cypriano E. S., Sodré, Jr. L., Kneib J.-P., Campusano L. E., 2004, ApJ, 613, 95
Ebeling H., Edge A. C., Mantz A., Barrett E., Henry J. P., Ma C. J., van Speybroeck L., 2010, MNRAS, 407, 83
Fedderke M. A., Kolb E. W., Wyman M., 2015, PRD, 91, 063505
Feldmann R., Mayer L., 2015, MNRAS, 446, 1939
Fiacconi D., Feldmann R., Mayer L., 2015, MNRAS, 446, 1957
Giovannini G., Tordi M., Feretti L., 1999, na, 4, 141
Girardi M., Mezzetti M., 2001, ApJ, 548, 79
Govoni F., Enßlin T. A., Feretti L., Giovannini G., 2001, A&A, 369, 441
Harrison E. R., 1967, Reviews of Modern Physics, 39, 862
Hezaveh Y., Dalal N., Holder G., Kisner T., Kuhlen M., Perreault Levasseur L., 2014, ArXiv e-prints: 1403.2720
Jauzac M. et al., 2015, MNRAS, 446, 4132
Jauzac M. et al., 2014, ArXiv e-prints: 1409.8663
Johnson T. L., Sharon K., Bayliss M. B., Gladders M. D., Coe D., Ebeling H., 2014, ApJ, 797, 48
Jullo E., Kneib J.-P., Limousin M., Elíasdóttir Á., Marshall P. J., Verdugo T., 2007, New Journal of Physics, 9, 447
Kelly P. L. et al., 2014, The Astronomer's Telegram, 6729, 1
Kelly P. L. et al., 2015, Science, 347, 1123
Kempner J. C., David L. P., 2004, MNRAS, 349, 385
Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102
Lam D., Broadhurst T., Diego J. M., Lim J., Coe D., Ford H. C., Zheng W., 2014, ApJ, 797, 98
Liesenborgs J., De Rijcke S., Dejonghe H., Bekaert P., 2007, MNRAS, 380, 1729
Liesenborgs J., De Rijcke S., Dejonghe H., Bekaert P., 2008, MNRAS, 389, 415

Liesenborgs J., De Rijcke S., Dejonghe H., Bekaert P., 2009, MNRAS, 397, 341
Limousin M. et al., 2012, A&A, 544, A71
Mann A. W., Ebeling H., 2012, MNRAS, 420, 2120
Marinacci F., Pakmor R., Springel V., 2014, MNRAS, 437, 1750
Martizzi D., Mohammed I., Teyssier R., Moore B., 2014, MNRAS, 440, 2290
Massey R. et al., 2015, MNRAS, 449, 3393
Medezinski E. et al., 2013, ApJ, 777, 43
Merten J. et al., 2011, MNRAS, 417, 333
Mohammed I., Liesenborgs J., Saha P., Williams L. L. R., 2014, MNRAS, 439, 2651
Mohammed I., Saha P., Liesenborgs J., 2015, pasj, 67, 21
Natarajan P., De Lucia G., Springel V., 2007, MNRAS, 376, 180
Natarajan P., Springel V., 2004, ApJL, 617, L13
Newman A. B., Treu T., Ellis R. S., Sand D. J., 2013a, ApJ, 765, 25
Newman A. B., Treu T., Ellis R. S., Sand D. J., Nipoti C., Richard J., Jullo E., 2013b, ApJ, 765, 24
Ogrean G. et al., 2015, ArXiv e-prints
Owers M. S., Randall S. W., Nulsen P. E. J., Couch W. J., David L. P., Kempner J. C., 2011, ApJ, 728, 27
Postman M. et al., 2012, ApJS, 199, 25
Press W. H., Schechter P., 1974, ApJ, 187, 425
Rau S., Vegetti S., White S. D. M., 2014, MNRAS, 443, 957
Reichardt C. L. et al., 2009, ApJ, 694, 1200
Richard J. et al., 2014, MNRAS, 444, 268
Richard J., Kneib J.-P., Limousin M., Edge A., Jullo E., 2010, MNRAS, 402, L44
Rodney S. A. et al., 2015, ArXiv e-prints
Schaye J. et al., 2015, MNRAS, 446, 521
Schrodinger E., 1939, Physica, 6, 899
Seljak U., 2000, MNRAS, 318, 203
Shan H., Qin B., Fort B., Tao C., Wu X.-P., Zhao H., 2010, MNRAS, 406, 1134
Sharon K., Johnson T. L., 2015, ApJL, 800, L26
Silk J., 1968, ApJ, 151, 459
Skillman S. W., Warren M. S., Turk M. J., Wechsler R. H., Holz D. E., Sutter P. M., 2014, ArXiv e-prints: 1407.2600
Smail I., Ellis R. S., Dressler A., Couch W. J., Oemler A., Sharples R. M., Butcher H., 1997, ApJ, 479, 70
Smith G. P. et al., 2009, ApJL, 707, L163
Soucail G., Mellier Y., Fort B., Mathez G., Cailloux M., 1987, The Messenger, 50, 5
Toshikawa J. et al., 2014, ApJ, 792, 15
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Vogelsberger M. et al., 2014, Nature, 509, 177
Williams L. L. R., Saha P., 2011, MNRAS, 415, 448
Zitrin A., Broadhurst T., 2009, ApJL, 703, L132
Zitrin A., Broadhurst T., Barkana R., Rephaeli Y., Benítez N., 2011, MNRAS, 410, 1939
Zitrin A. et al., 2015, ApJ, 801, 44
Zitrin A. et al., 2013, ApJL, 762, L30

7

Conclusions

In this PhD dissertation, I (along with other collaborators on different projects) tried to build a better understanding of the distribution of matter in the Universe and its clustering properties, which I studied in two very different regimes: gravitationally bound systems like cluster of galaxies, and the large scale structure of the Universe. While working on various projects on this topic, I successfully completed five scientific publications:

- The very first project is a mass modelling problem in strong lensing cluster [49]. Here we used a publicly available code GRALE, modified it for an optimum solution and resolution, to put tight constraints on the mass distribution of few lensing clusters. By studying the mass and light distribution of these clusters, we imply that it is possible that the dark-matter has a finite self-interaction cross section and shows important signatures in the central parts of clusters.

- In a similar project, we show that the time delay information is very useful in order to put additional constraints on the central region of the clusters which can't be resolved directly, especially when the steepness degeneracy is broken by the presence of background sources at different redshift [51].

- We also presented a (semi) analytic model for the matter power spectrum, which is computationally inexpensive and computes the power spectrum to a percent level accuracy up to k $\sim 1\ h^{-1}$Mpc [52]. The motivation of this model is the halo model and we also derived a simple form of the covariance matrix. We proposed a way to marginalise over baryonic effects.

- In another project, we used the halo model in order to directly model the effects of baryons on the matter power spectrum and to the extension on the weak lensing shear power spectrum [50]. The effects are small at comparatively large scales but as the next generation surveys are expected to measure these quantities to very small scales, the baryonic effects are important to take into

account. If not, it will add biases to the cosmological parameters up to 10 sigma.

- Finally, we produced free-form mass models for six Hubble Frontier Field clusters. Their mass distribution shows elongation, multiple-cores, and many sub-structures indicating a recent major merger. We measured the power spectrum of the mass distribution to quantify the sub-structures, and compared them to the simulated clusters.

7.0.2. A unified picture

If we try to draw a bigger picture from all the projects above, it states that it is very important to model baryonic physics in order to understand the clustering processes, galaxy formation, and to do cosmology with future generation surveys.

Gravitational lensing is one ideal tool to do accomplish this. Strong lensing is very useful in studying the individual systems. Weak lensing shear measurements on larger area in the sky is the most promising tool to do cosmology under control systematics. One of the biggest source of systematics is again the baryonic effects at small scales, and there are various ways to handle them. But they can't be ignored. Precision cosmology and detailed information about the individual systems is needed in order to gain full understanding of the galaxy formation processes and evolution of the Universe.